ARJ

Defense Acquisition Research Journal

A Publication of the Defense Acquisition University

Thinking Small in Order to THINK BIG

Defense Acquisition University

January 2017 Vol. 24 No. 1 | ISSUE 80

Article List

Survey of Small Business Barriers to Department of Defense Contracts
Ronnie Schilling, Thomas A. Mazzuchi, and Shahram Sarkani

Using Heuristics for Supportability Analysis of Adaptive Weapon Systems in Combat
Samuel H. Amber

The Threat Detection System That Cried Wolf: Reconciling Developers with Operators
Shelley M. Cazares

Increasing Army Supply Chain Performance Using an Integrated End-to-End Metrics System
Fan T. Tseng, Laird Burns, James T. Simpson, and David Berkowitz

Scandal and Tragedy? Or Acquisition Lessons Relearned by the F-35 Program
Col Roger Witek, USAF (Ret.)

Online-only Article
Federally Mandated Furloughs: The Effect on Organizational Commitment and Organizational Citizenship Behavior
Robert L. Shepherd

ARJ Extra

The Defense Acquisition Professional Reading List
Rational Action: The Sciences of Policy in Britain and America, 1940–1960
Written by William Thomas
Reviewed by Petros Boutselis

Defense Acquisition Research Journal
A Publication of the Defense Acquisition University

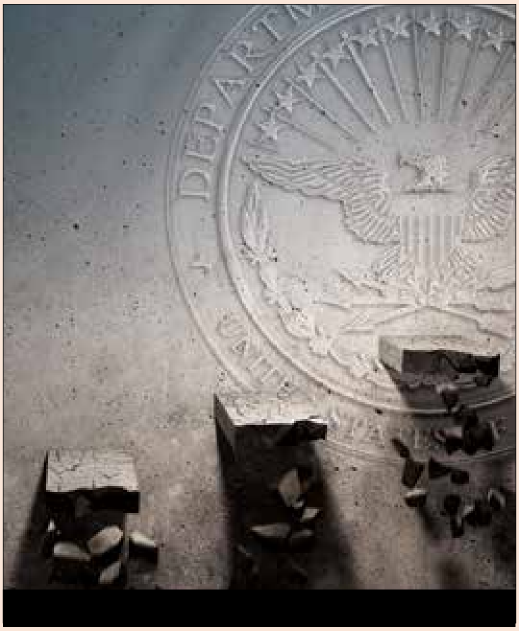
ISSN 2156-8391 (print) ISSN 2156-8405 (online)
DOI: http://dx.doi.org/10.22594/dau.012017-80.24.01

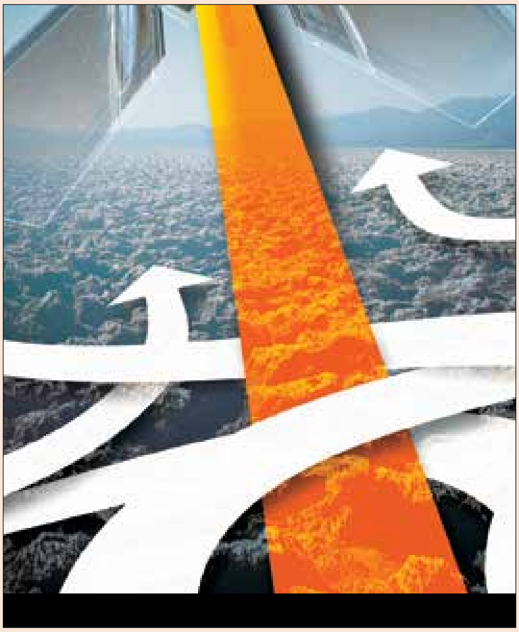
The *Defense Acquisition Research Journal*, formerly the *Defense Acquisition Review Journal*, is published quarterly by the Defense Acquisition University (DAU) Press and is an official publication of the Department of Defense. Postage is paid at the U.S. Postal facility, Fort Belvoir, VA, and at additional U.S. Postal facilities. Postmaster, send address changes to: Editor, *Defense Acquisition Research Journal*, DAU Press, 9820 Belvoir Road, Suite 3, Fort Belvoir, VA 22060-5565. The journal-level DOI is: http://dx.doi.org/10.22594/dauARJ.issn.2156-8391. Some photos appearing in this publication may be digitally enhanced.

Defense Acquisition Research Journal

A Publication of the Defense Acquisition University

CONTENTS | Featured Research



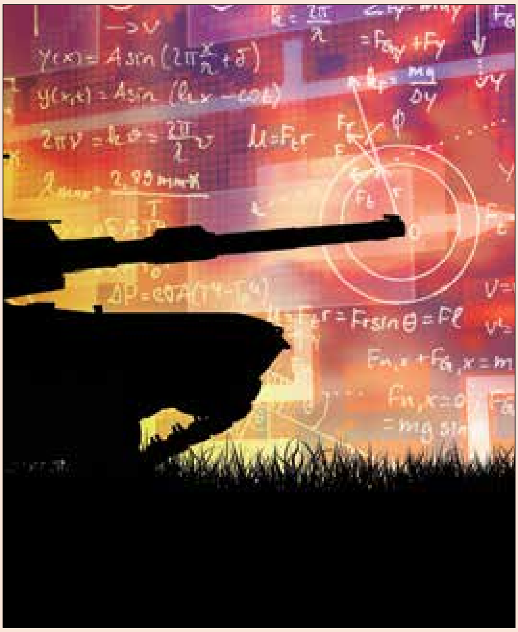


p. **42**

The Threat Detection System That Cried Wolf: Reconciling Developers with Operators

Shelley M. Cazares

Threat detection systems that perform well in testing can "cry wolf" during operation, generating many false alarms. The author posits that program managers can still use these systems as part of a tiered system that, overall, exhibits better performance than each individual system alone.



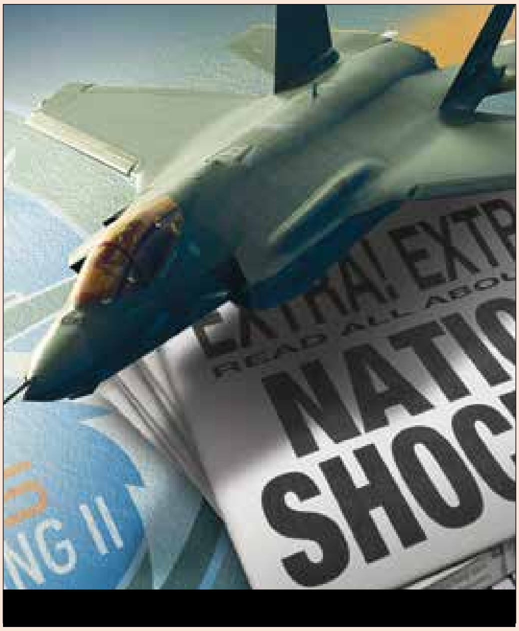
p. **66**

Increasing Army Supply Chain Performance Using an Integrated End-to-End Metrics System

Fan T. Tseng, Laird Burns, James T. Simpson, and David Berkowitz

The authors provide insights into a research project to develop an end-to-end performance metrics system framework for the U.S. Army Materiel Command. The performance metrics system framework enables reporting of top-level strategic metrics with full drill-down capability, as illustrated by the development of the Customer Wait Time metric that was adopted by Joint Supply Chain Architecture.

CONTENTS | Featured Research



http://www.dau.mil/publications/DefenseARJ/default.aspx

CONTENTS | Featured Research

FROM THE CHAIRMAN AND EXECUTIVE EDITOR

Dr. Larrie D. Ferreiro

The theme for this edition of *Defense Acquisition Research Journal* is "Thinking Small in Order to Think Big," as many of the articles drill down into the important details of processes and procedures in order to develop larger lessons for defense acquisition. The first article, "Survey of Small Business Barriers to Department of Defense Contracts," by Ronnie Schilling, Thomas A. Mazzuchi, and Shahram Sarkani, examines the factors that small businesses see as inhibiting them from pursuing defense contracts, which Better Buying Power specifically attempts to encourage. They found that lack of communications and long timelines for approvals and decisions were some of the most important reasons cited by small business leaders. The next article, "Using Heuristics for Supportability Analysis of Adaptive Weapon Systems in Combat," by Samuel H. Amber argues that, given the difficulty of obtaining supportability data on deployed weapon systems that often have been modified for combat, incorporating heuristics as an alternative field data source in the decision matrix can improve the development of supportability requirements.

Shelley M. Cazares, in "The Threat Detection System That Cried Wolf: Reconciling Developers with Operators," points out that some threat detection systems, which perform well in testing, can generate many false alarms or "cry wolf" in operation. One way to mitigate this problem may be to use these systems as part of a tiered system

that, overall, exhibits better performance than each individual system alone. Next, "Increasing Army Supply Chain Performance Using an Integrated End-to-End Metrics System" by Fan T. Tseng, Laird Burns, James T. Simpson, and David Berkowitz describes a process of pulling information from multiple data systems into a single integrated end-to-end performance metrics system that enables a high-level dashboard overview as well as full drill-down capabilities to source-level data and documents. Finally, thinking big rather than thinking small, Roger Witek addresses one of the largest defense acquisition programs in history in "Scandal and Tragedy? Or Acquisition Lessons Relearned by the F-35 Program," in which he compares the Joint Strike Fighter (JSF) program with previous joint aircraft acquisition programs, and mines observations from think tanks and policy experts in order to weave together a compendium of lessons learned, lessons relearned, and recommendations for future acquisition strategies.

The article "Federally Mandated Furloughs: The Effect on Organizational Commitment and Organizational Citizenship Behavior" by Robert L. Shepherd is available in full in the online edition of the *Defense Acquisition Research Journal*, and abstracted in the printed version. It identifies the negative effects associated with furloughs, such as decreases in morale, lowered productivity rates, and increases in employee turnover so that managers can more effectively address them.

The featured book in this issue's Defense Acquisition Professional Reading List is *Rational Action: The Sciences of Policy in Britain and America, 1940–1960* by William Thomas, reviewed by Petros Boutselis of the Centre for Defence Acquisition at Cranfield University UK.

DAU CENTER FOR DEFENSE ACQUISITION

RESEARCH AGENDA 2017–2018

This Research Agenda is intended to make researchers aware of the topics that are, or should be, of particular concern to the broader defense acquisition community within the federal government, academia, and defense industrial sectors. The center compiles the agenda annually, using inputs from subject matter experts across those sectors. Topics are periodically vetted and updated by the DAU Center's Research Advisory Board to ensure they address current areas of strategic interest.

The purpose of conducting research in these areas is to provide solid, empirically based findings to create a broad body of knowledge that can inform the development of policies, procedures, and processes in defense acquisition, and to help shape the thought leadership for the acquisition community. Most of these research topics were selected to support the DoD's Better Buying Power Initiative (see http://bbp.dau.mil). Some questions may cross topics and thus appear in multiple research areas.

Potential researchers are encouraged to contact the DAU Director of Research (research@dau.mil) to suggest additional research questions and topics. They are also encouraged to contact the listed Points of Contact (POC), who may be able to provide general guidance as to current areas of interest, potential sources of information, etc.

Competition POCs

- John Cannaday, DAU: john.cannaday@dau.mil
- Salvatore Cianci, DAU: salvatore.cianci@dau.mil
- Frank Kenlon (global market outreach), DAU: frank.kenlon@dau.mil

Measuring the Effects of Competition

- What means are there (or can be developed) to measure the effect on defense acquisition costs of maintaining the defense industrial base in various sectors?
- What means are there (or can be developed) of measuring the effect of utilizing defense industrial infrastructure for commercial manufacture, and in particular, in growth industries? In other words, can we measure the effect of using defense manufacturing to expand the buyer base?
- What means are there (or can be developed) to determine the degree of openness that exists in competitive awards?
- What are the different effects of the two best value source selection processes (trade-off vs. lowest price technically acceptable) on program cost, schedule, and performance?

Strategic Competition

- Is there evidence that competition between system portfolios is an effective means of controlling price and costs?
- Does lack of competition automatically mean higher prices? For example, is there evidence that sole source can result in lower overall administrative costs at both the government and industry levels, to the effect of lowering total costs?
- What are the long-term historical trends for competition guidance and practice in defense acquisition policies and practices?

- To what extent are contracts being awarded noncompetitively by congressional mandate for policy interest reasons? What is the effect on contract price and performance?

- What means are there (or can be developed) to determine the degree to which competitive program costs are negatively affected by laws and regulations such as the Berry Amendment, Buy American Act, etc.?

- The DoD should have enormous buying power and the ability to influence supplier prices. Is this the case? Examine the potential change in cost performance due to greater centralization of buying organizations or strategies.

Effects of Industrial Base

- What are the effects on program cost, schedule, and performance of having more or fewer competitors? What measures are there to determine these effects?

- What means are there (or can be developed) to measure the breadth and depth of the industrial base in various sectors that go beyond simple head-count of providers?

- Has change in the defense industrial base resulted in actual change in output? How is that measured?

Competitive Contracting

- Commercial industry often cultivates long-term, exclusive (noncompetitive) supply chain relationships. Does this model have any application to defense acquisition? Under what conditions/circumstances?

- What is the effect on program cost, schedule, and performance of awards based on varying levels of competition: (a) “Effective” competition (two or more offers); (b) “Ineffective” competition (only one offer received in response to competitive solicitation); (c) split awards vs. winner take all; and (d) sole source.

Improve DoD Outreach for Technology and Products from Global Markets

- How have militaries in the past benefited from global technology development?
- How/why have militaries missed the largest technological advances?
- What are the key areas that require the DoD's focus and attention in the coming years to maintain or enhance the technological advantage of its weapon systems and equipment?
- What types of efforts should the DoD consider pursuing to increase the breadth and depth of technology push efforts in DoD acquisition programs?
- How effectively are the DoD's global science and technology investments transitioned into DoD acquisition programs?
- Are the DoD's applied research and development (i.e., acquisition program) investments effectively pursuing and using sources of global technology to affordably meet current and future DoD acquisition program requirements? If not, what steps could the DoD take to improve its performance in these two areas?
- What are the strengths and weaknesses of the DoD's global defense technology investment approach as compared to the approaches used by other nations?
- What are the strengths and weaknesses of the DoD's global defense technology investment approach as compared to the approaches used by the private sector—both domestic and foreign entities (companies, universities, private-public partnerships, think tanks, etc.)?
- How does the DoD currently assess the relative benefits and risks associated with global versus U.S. sourcing of key technologies used in DoD acquisition programs? How could the DoD improve its policies and procedures in this area to enhance the benefits of global technology sourcing while minimizing potential risks?

- How could current DoD/U.S. Technology Security and Foreign Disclosure (TSFD) decision-making policies and processes be improved to help the DoD better balance the benefits and risks associated with potential global sourcing of key technologies used in current and future DoD acquisition programs?

- How do DoD primes and key subcontractors currently assess the relative benefits and risks associated with global versus U.S. sourcing of key technologies used in DoD acquisition programs? How could they improve their contractor policies and procedures in this area to enhance the benefits of global technology sourcing while minimizing potential risks?

- How could current U.S. Export Control System decision-making policies and processes be improved to help the DoD better balance the benefits and risks associated with potential global sourcing of key technologies used in current and future DoD acquisition programs?

Comparative Studies

- Compare the industrial policies of military acquisition in different nations and the policy impacts on acquisition outcomes.

- Compare the cost and contract performance of highly regulated public utilities with nonregulated "natural monopolies," e.g., military satellites, warship building, etc.

- Compare contracting/competition practices between the DoD and complex, custom-built commercial products (e.g., offshore oil platforms).

- Compare program cost performance in various market sectors: highly competitive (multiple offerors), limited (two or three offerors), monopoly?

- Compare the cost and contract performance of military acquisition programs in nations having single "purple" acquisition organizations with those having Service-level acquisition agencies.

ISSUE 80
JANUARY 2017
VOL. 24 NO. 1

We’re on the Web at:
http://www.dau.mil/publications/DefenseARJ/default.aspx
and
http://dau.dodlive.mil

Survey of **SMALL BUSINESS BARRIERS** to DEPARTMENT of **DEFENSE CONTRACTS**

Ronnie Schilling, Thomas A. Mazzuchi, and Shahram Sarkani

A key tenet of the Better Buying Power initiatives is to increase small business participation in Department of Defense contracting. The department has had mixed success in retaining small businesses and meeting small business contracting goals. Results of a survey given to 681 small business leaders show many factors commonly exist that prevent small businesses from pursuing defense contracts. Some factors are more common than others, with the most cited factors related to a lack of communication from government leads or to the government taking too long to give approvals and make decisions. Statistical evidence also supports the perceptions, of smaller and newer small businesses, that the defense business is more challenging for them than for their larger and more experienced competitors. However, this turned out to be the case for only a subset of the factors we explored.

DOI: *http://dx.doi.org/10.22594/dau.16-752.24.01*
Keywords: *challenges, competition, statistical analysis*

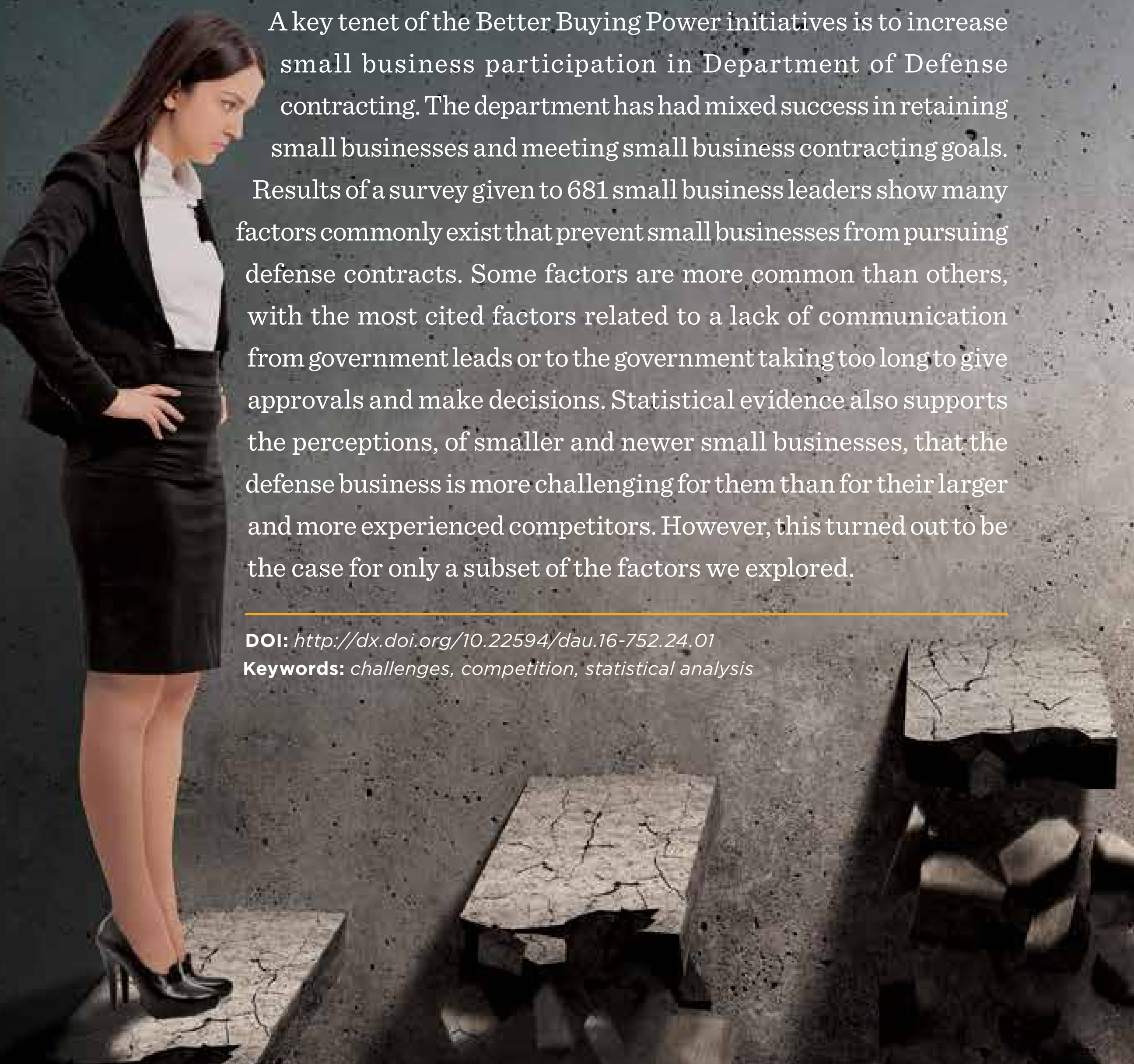


Image designed by Diane Fleischer

Creating more than half of this country's gross domestic product and seven out of every 10 new jobs (Graves, n.d., p. 1), small business holds a place of importance in the U.S. economy that cannot be overstated. Small businesses are also a key driver of innovation, producing on average 13 times more patents per employee than large firms (Mielach, 2012, para 4).

Small businesses are also an important contributor in defense acquisition, providing innovation, competition, and services at great value (Simmers, 2011). The Department of Defense (DoD) has recognized that it must attract and retain small businesses in order to continue to create and maintain world-class weapon systems. Department leadership has pushed for increased small business roles and opportunities through the Better Buying Power initiatives, seeking to provide "the maximum practicable opportunity" for small businesses (Kendall, 2015, p. 25). The department has also set a small business prime contracting goal, which has ranged from 21 to 23% of all contract obligations per year from 2011 through 2016.

The DoD has had mixed success in meeting its small business prime contracting goal. It had missed this goal for 7 years in a row before finally meeting it in 2014 (Serbu, 2015, para 1).

The DoD has had mixed success in meeting its small business prime contracting goal. It had missed this goal for 7 years in a row before finally meeting it in 2014 (Serbu, 2015, para 1). Some attribute this to the nature of what the department buys: aircraft, tanks, and large weapon systems often beyond the capabilities of small businesses (Lee, 2012). However, critics of the Pentagon's small business policies claim it is too simple and self-centered (Chandler, 2014), focusing on a single contract obligation goal that does not take into account the needs and challenges from the perspective of the small businesses themselves. Businesses will respond to a government request for proposal (RFP) only if it presents a satisfactory business case, something the government often overlooks (Chandler, 2014). Attracting small businesses to conduct business with DoD has been a challenge. An analysis of small businesses in the Central Contractor Registration database showed that of the small businesses conducting business with DoD from 1997 to 2007, 44% had stopped conducting business with DoD in 2008 (Moore et al., 2008, p. 85). Only 1.9% of the businesses had continuous

contracts with DoD over the 11 years. Further, 46% of small businesses that received contracts during this period received only one contract, which was valid for a year or less, possibly indicating an unwillingness of businesses to pursue additional opportunities after their initial contract was complete (Moore et al., 2008, p. 72).

The challenges facing businesses in defense contracting have been previously well documented. The Lexington Institute recently published an article calling the DoD a “difficult buyer” that discourages businesses from submitting bids on defense contracts and proposing that industry and government take more of a partnership approach to contracting (Chandler, 2014). The Government Accountability Office (GAO) has also published multiple reports claiming the DoD needs to improve reporting and tracking of small business financial obligations (Neumann, 2015; Shear, 2011; Woods, 2013). Most recently, the challenges related to small businesses have gotten the attention of Congress, and the House Armed Services Committee (HASC) directed a panel to investigate challenges currently affecting the defense industry base.

Factors Affecting Small Business Participation in DoD Contracts

We conducted a comprehensive literature review to better understand the factors preventing small businesses from desiring to conduct business with DoD. This review included previous panel interviews of business owners, including the 2012 HASC study; opinionated articles; case studies; and GAO reports. The purpose of this review was not to determine whether the perceptions presented in these articles were valid, but simply to identify what factors exist. We found 26 factors that are potential barriers preventing businesses

from pursuing an initial contract or follow-on contracts with DoD. The factors we found fit into three categories: the contract solicitation and award process, contract requirements, and contract execution. A summary of the barriers we found in each category is provided below.

The Contract Solicitation and Award Process

The contract solicitation documentation and the metrics for evaluating proposals were cited as being too complex and difficult to understand by business leaders (HASC, 2012). A lack of communication during the solicitation process was also cited as an issue, with businesses struggling to get answers to their questions on solicitation documentation and desiring more feedback on proposals that did not win a contract (Maser & Thompson, 2013). Given the complexity of the solicitation documentation, the amount of paperwork required to submit proposals and insufficient time to develop

proposals were also cited as barriers (Krieger, 2015). Previous studies also found that delays in making contract awards have created financial hardships on small businesses (Chaplain, 2010).

> ***The literature review revealed that many small businesses continue to struggle with defense business after winning contracts.***

Contract Requirements

Defense contracts have unique requirements for placing a bid that are not found in the commercial world. While many of these requirements serve a good purpose—often to protect the government—they were also cited to be barriers to defense business in the articles we reviewed. Contract requirements that were cited as being issues in previous studies include surety bonds, government cost accounting standards, export control regulations (International Traffic in Arms), and a past performance rating on previous government contracts (HASC, 2012). Regulations, both the complexity and the number of them, were also cited as an issue (Friar, 2015). Protecting proprietary data was another area of concern for businesses in previous case studies (Chaplain, 2010). Negotiation of data rights has been getting an increased amount of attention in the last few years, both from industry and the government (Erwin, 2014). Another issue with contract requirements cited in previous studies is technical requirements being written too narrowly or seemingly catering to a particular vendor (Chaplain, 2010), thus preventing other businesses from reasonably competing for these solicitations. Contractor profits have long been a source of contention and were also cited as a barrier to defense contracts, because many businesses can make higher margins in the commercial sector (Chandler, 2014).

Contract Execution

The literature review revealed that many small businesses continue to struggle with defense business after winning contracts. Cited barriers associated with contract execution include issues getting payments for completed work or payments taking too long, challenges contacting government leads or government leads not being helpful when contacted, contract modifications taking too long, and approvals taking too long (Krieger, 2015; Mills, 2010).

Research Hypothesis Methodology

Based on our literature review, there is no question that barriers exist and that they prevent small businesses from desiring to pursue defense contracts. What is unknown is how common these barriers are. Are they widespread issues or isolated incidents? We wanted to gain more insight into these barriers and how they affect small business participation in defense contracts. Specifically, we wanted to determine the following:

- How common are the previously identified barriers to defense business? The prior works cited above are either opinionated articles or are studies conducted with a handful of small businesses. A recent large-scale study with quantitative data on these barriers does not exist.

- How do these barriers vary by industry or business type? DoD conducts business in numerous areas, and previous studies have argued that a rigid "one size fits all" contracting approach may not be appropriate (Blakey, 2011, p. 4).

- Do new small businesses perceive any of these potential barriers to be more of an issue than those that have extensive experience conducting defense business? Studies have shown that "nontraditional" small businesses with little to no government experience struggle to compete for defense contracts (Cox, Moore, & Grammich, 2014). Many of these nontraditional small businesses have extensive commercial experience and/or new and innovative technologies that the department could use (Freedberg, 2014).

- Do smaller small businesses perceive any of these barriers to be more of an issue than those that are larger but still qualify as a "small business"? The qualifications to be considered a small business vary by industry, but are on the order of 500 to 1,500 employees. Previous studies have cited these qualifications as too large (Bail, 2010), stating that smaller small businesses with tens of employees cannot compete for small business set-asides against those with hundreds or thousands of employees.

To gain further insight into our research hypothesis, we developed and administered a survey for small businesses. A large-scale survey was administered to small businesses that had formerly conducted business with or are currently conducting business with the DoD. A mailing list was

developed from contract award data and small business registries. An electronic mail invitation was sent directly to the business CEO, president, or business development lead if the information was available (this information was available for 78% of the invitations that were sent). Otherwise, the invitation was sent to a publicly known address accompanied by a request that it be forwarded to the appropriate point of contact. The survey was conducted from September to November 2015, and we received 681 responses from small businesses.

The first portion of the survey consisted of collecting demographic information about the business: primary line(s) of business, number of employees, and business history with DoD. The participants were then given a list of 26 factors and asked to rate the importance of each factor in their decision not to pursue additional defense business opportunities. The 26 factors were selected based on our literature review of previous articles on the subject. Each of the factors was rated using a 5-point Likert scale ranging from "not a factor" to "very important factor." The respondents were also given

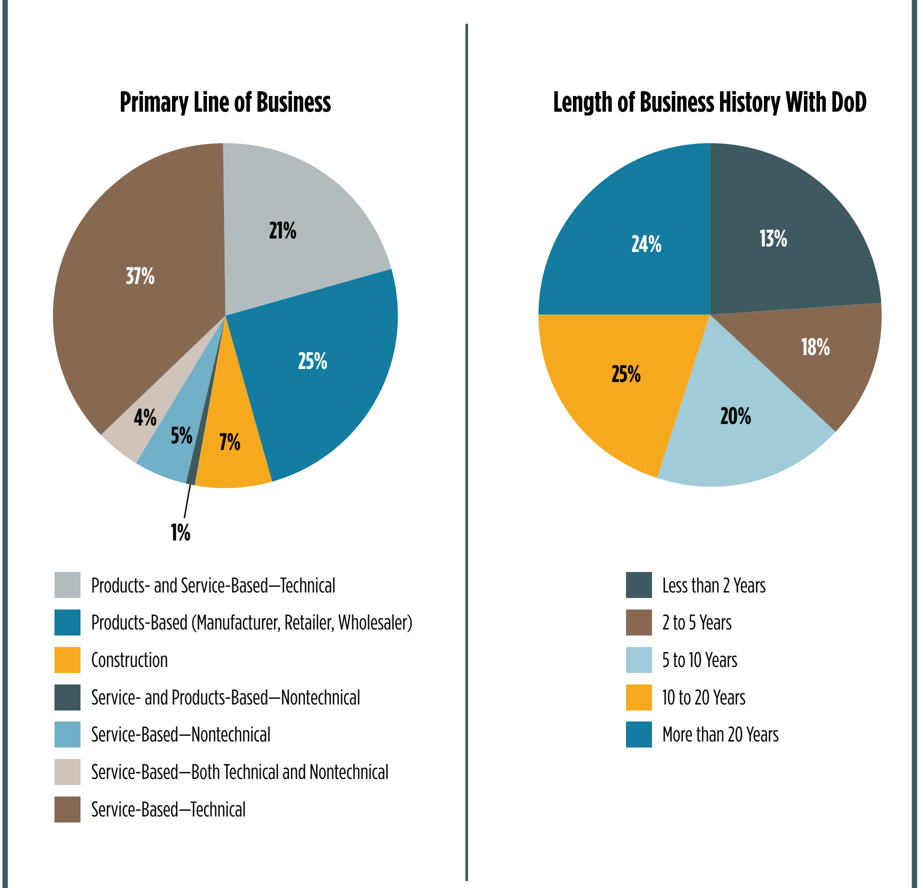
an option to select “don’t know or no opinion” for each factor; for analysis purposes, we discarded both these responses and responses that were left unanswered. Participants were also given three open-ended qualitative questions to express any other potential barriers not considered in the Likert-scale questions, to provide examples, and to offer suggestions on how to make defense business more attractive.

Demographics

Survey respondents included businesses representing the major areas in which DoD conducts business. Respondents included “technical” service-based businesses such as small business innovation research (SBIR) participants and engineering support firms, “nontechnical” service-based businesses providing services DoD uses every day (such as janitorial and grounds maintenance), product-based businesses such as manufacturers

and retailers, and construction businesses. Many of the respondents conduct business with DoD in multiple areas. Respondents' length of experience conducting business with DoD spanned a wide range, from businesses relatively new to the industry (1–2 years) to ones conducting business with DoD for more than 20 years (Figure 1).

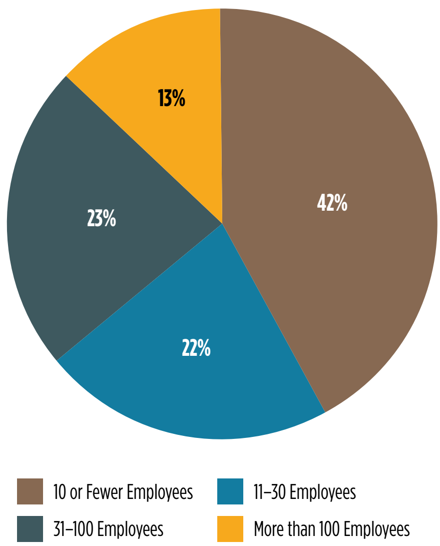
FIGURE 1. SURVEY RESPONDENTS: PRIMARY LINE OF BUSINESS AND LENGTH OF BUSINESS HISTORY



Respondents also included small businesses having a fairly wide range in size. Many respondents had only a handful of employees working for their business; 42% had 10 or fewer, and 64% had 30 or fewer. The large majority—87%—had fewer than 100 employees. We also received responses from

the "larger" businesses with several hundred employees that do still qualify for small business contracts, making up 13% of the total survey respondents (Figure 2).

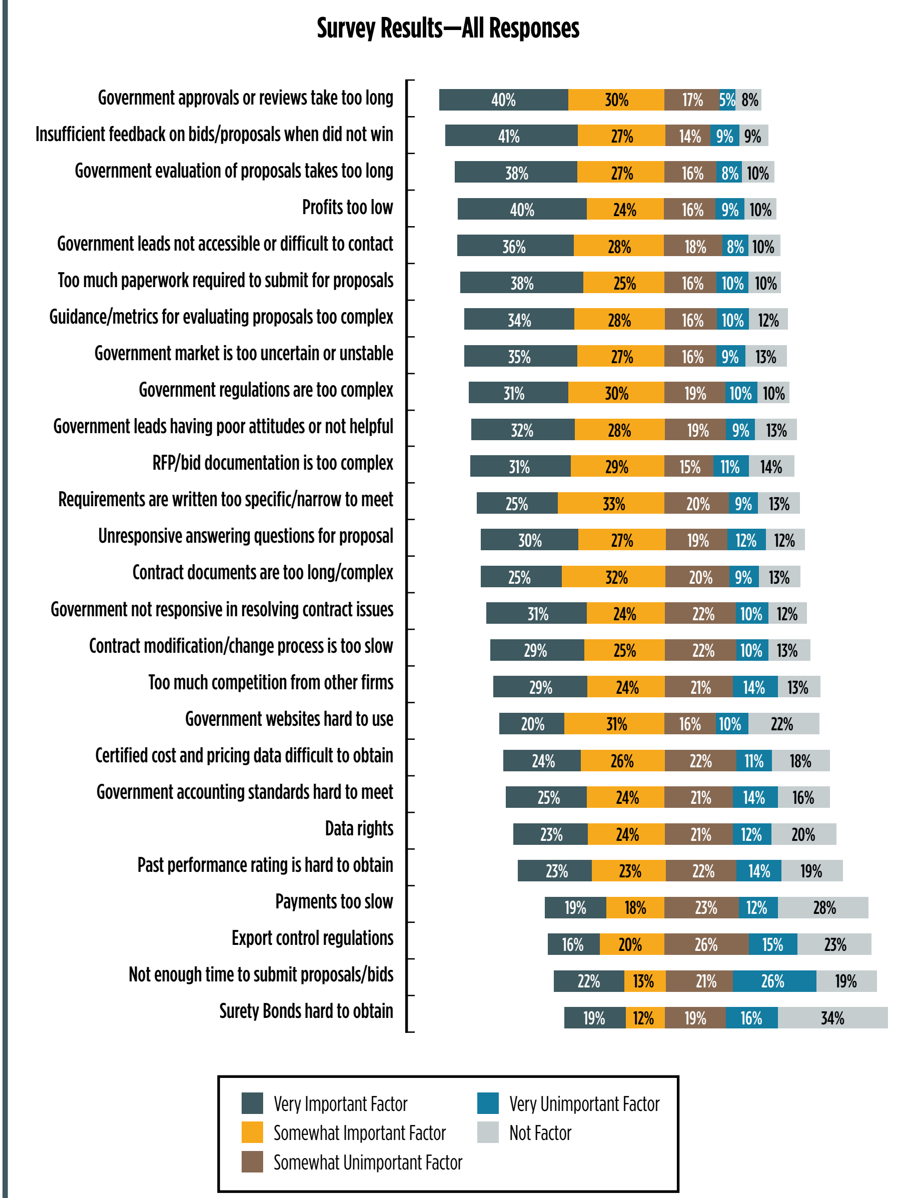
FIGURE 2. SURVEY RESPONDENTS: NUMBER OF EMPLOYEES



Quantitative Survey Results

The survey results show that of the 26 potential barriers surveyed, most were perceived to exist to a reasonably common extent; 19 of them were rated as a "somewhat important" or "very important" barrier by more than half of the respondents in their decision not to pursue additional business opportunities with the DoD. A smaller, although fairly significant, proportion of the respondents cited the barriers as being a "non-issue" or only "very minor," ranging from 13% to half of the respondents. The full results are shown in Figure 3.

FIGURE 3. SURVEY RESULTS



The barriers in Figure 3 are ranked from top to bottom in the order they were cited as being “very important” or “somewhat important,” the most to the least. Four of the five top-ranked barriers—approvals taking too long, insufficient feedback on proposals, proposal evaluations taking too long, and inaccessible government leads—are directly related to a lack of communication from government leads or to the government taking too long to make approvals and decisions. Sixty-five to 70% of respondents rated these as either “somewhat important” or “very important.” Waiting for approvals was the top-ranked factor, with many respondents giving examples of waiting for approvals during both the contract award process and during contract execution. One respondent provided an example in which it took months to get approved to place a bid for parts, only to get approval after the solicitation had closed. Contract award times on the order of 6 months to a year for simple contracts were commonly cited by the respondents. One respondent gave an example of choosing to “no-bid” a phase II SBIR contract after winning a phase I contract, explaining that they could not afford to keep the company’s principal investigator on the payroll for the amount of time the government takes to award the contract. One respondent questioned how the government can take excessive times to award contracts and give minimal time for businesses to prepare proposals and suggested the government be held to the same timeframe as they are. The final barrier ranked in the top five was low profits, long a source of contention, which was cited as the fourth most important barrier, with 64% of respondents naming it as an important issue preventing them from pursuing additional defense business opportunities. Many respondents referred to profit levels decreasing on government contracts, often to “single digit levels.” Respondents said that profit levels of 10 to 15% would normally be acceptable. However,

defense contracts require extensive overhead to develop proposals and meet other contract requirements, such as accounting standards, surety bonds, and security clearances. This “excessive” overhead further eats into profits and makes profit levels of 10 to 15% unattractive and profit levels less than this unsustainable.

> ***Numerous respondents also felt that government regulations favor larger businesses, as small businesses have to comply with the same regulations, but cannot hire proper personnel to understand them.***

Sixty to 63% of respondents ranked three of the remaining barriers, among the 10 most cited, as “somewhat important” or “very important.” These are associated with the large amount of effort and specialized knowledge required by the defense acquisition process: too much paperwork required to submit proposals (ranked sixth), guidance/metrics for evaluating proposals being too complex (ranked seventh), and government regulations being too complex (ranked ninth). Many respondents cited the need to hire subject matter experts to fully understand government solicitations and regulations. Numerous respondents also expressed that government regulations favor larger businesses, as small businesses have to comply with the same regulations, but cannot hire proper personnel to understand them. Respondents cited costs for the business and planning effort to develop proposals on the order of tens to hundreds of thousands of dollars. This is significantly higher than commercial proposal costs, leading them to pursue commercial contracts in place of defense contracts. The government market being too unstable was ranked eighth, with many of the respondents noting that the recent instability of budgets and contract awards has reduced their ability to plan and conduct business development, also further eating into profits. Government personnel not being helpful when contacted was ranked 10th.

Following the top 10 cited barriers, the factors are largely focused on the contract-solicitation-and-award process and issues with contract documents. For the contract-solicitation-and-award process: proposal documentation being too complex was ranked 11th, requirements in the

solicitation being written too narrowly to be reasonably met was ranked 12th, and the government not being responsive in answering questions during the solicitation period was ranked 13th. Many respondents noted that they had regularly noticed solicitations put out for competition that were seemingly tailored to a particular vendor, as the solicitation had very specific requirements that no other vendor could reasonably meet. For issues with contract documents: contract documents being too long and complex was ranked 14th, the government not being responsive in resolving contract issues was ranked 15th, and the contract modification process being too slow was ranked 16th. One respondent noted that the slow pace of contract and requirement modifications limited the number of innovations they could provide to the government.

Contract requirements unique to defense contracts were among the lowest ranked barriers. Certified cost data, government accounting standards, data rights, past performance ratings, export control regulations, and surety bonds were ranked 19th, 20th, 21st, 22nd, 24th, and 26th, respectively. In general, a minority of the respondents—31 to 50%—rated these factors as "somewhat important" or "very important." This indicates that although these were cited in previous studies and in this study, they are less widespread as barriers for small businesses than the other barriers already mentioned.

Differences by Industry

We divided the survey responses into four independent groups, based on primary line of business: construction, technical service-based, nontechnical service-based, and product-based. For this portion of the analysis, respondents that conducted business in multiple areas were discarded.

The Mann-Whitney and Kruskal-Wallis statistical tests are commonly used to determine whether population medians are equal or differ among independent groups. The tests use a ranking methodology in which all responses for a factor are ranked from largest to smallest and then compare the mean ranking for each independent group. Table 1 provides the factors that differed enough between industries to be statistically significant, based on pairwise Mann-Whitney analysis. The z values for the Kruskal-Wallis test are also shown. The z value helps to interpret how the average rank for each group compares to the ranks from all groups. A negative z value indicates that the particular group perceived the barrier to be less important when compared to all groups, and a higher z value indicates that the particular group perceived the barrier to be more important when compared to all groups. As the absolute value of the z value gets larger, the further away a

particular group's average rank gets from the overall average rank. A greater absolute value indicates that the group perceived the barrier to be more or less of a barrier to a greater extent when compared to a lesser absolute value. Based on these tests, we found that the majority of the factors—21 of the 26—presented no statistically significant results. The five that did present statistically significant results are shown in Table 1.

TABLE 1. STATISTICALLY SIGNIFICANT DIFFERENCES BETWEEN INDUSTRIES

Barrier	Result	Kruskal-Wallis *z* value
Government websites hard to use	Product-based small businesses perceive this to be MORE of a barrier than the other three industries do.	Product: 2.80 Technical: -1.91 Nontechnical: -1.14 Construction: -0.34
Data rights: government requesting proprietary information	Technical service-based and product-based small businesses perceive this to be MORE of a barrier than nontechnical service-based and construction small businesses.	Product: 2.43 Technical: -0.27 Nontechnical: -1.70 Construction: -2.33
Export control regulations	Technical service-based and product-based small businesses perceive this to be MORE of a barrier than nontechnical service-based and construction small businesses.	Product: 0.72 Technical: 1.67 Nontechnical: -2.26 Construction: -2.72
Payment issues	Product-based small businesses perceive this to be LESS of a barrier than the other three industries do.	Product: -3.36 Technical: 2.30 Nontechnical: 0.96 Construction: 0.69
Surety bonds being difficult to obtain	Technical service-based small businesses perceive this to be LESS of a barrier than the other three industries do.	Product: 1.59 Technical: -3.76 Non-Technical: 1.17 Construction: 2.84

Based on the Mann-Whitney results, technical service-based and product-based small businesses found data rights and export control regulations to be more of a barrier compared to those that provide nontechnical products and services to the DoD. This doesn't come as much of a surprise, as these two items generally do not apply to the other two industries. Technical service-based small businesses have a relatively easier time obtaining surety bonds for their contracts. Small businesses that produce products for DoD have a relatively easier time getting paid for their work but have a harder time finding opportunities on government websites.

Experience as a Factor

We used the Spearman's rho correlation coefficient as a preliminary analysis of the relationship between each of the factors and length of business history with DoD. The coefficient measured the relationship between each of the factors and groups of businesses with various levels of experience (Figure 1), both of which are ordinal data. Fourteen of the factors presented correlation coefficients that were statistically significant. Of these 14 coefficients, all but one were negative, ranging from -.099 to -.263, indicating that the factors become less of a barrier as the business's level of experience increases. The one factor that provided a statistically significant positive trend was "government market being too uncertain or unstable," with a correlation coefficient of .084.

We broke the responses into independent groups by level of experience and again used the Mann-Whitney and Kruskal-Wallis statistical tests to look for differences. Ten of the 26 factors did present statistical differences between the groups based on the Mann-Whitney pairwise analysis. These 10 factors and associated Kruskal-Wallis z values are shown in Table 2.

TABLE 2. STATISTICALLY SIGNIFICANT DIFFERENCES BETWEEN LEVELS OF DoD EXPERIENCE

Barrier	Result	Kruskal-Wallis *z* value
Government websites hard to use	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than the groups of 10–20 and more than 20 years of experience.	< 2 years: 2.06
		2-5 years: -0.35
		5-10 years: 1.33
		10-20 years: -1.53
		> 20 years: -0.97
Government not providing enough time to submit proposals/bids	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than the groups of businesses with 10–20 and more than 20 years of experience.	< 2 years: 2.12
		2-5 years: -0.65
		5-10 years: 1.43
		10-20 years: -1.09
		> 20 years: -1.30
Insufficient feedback on bids/proposals when failing to win contract	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than the groups of 10–20 and more than 20 years of experience.	< 2 years: 2.25
		2-5 years: 0.82
		5-10 years: 0.46
		10-20 years: -1.11
		> 20 years: -1.94
Request for proposal/quotation/bid documentation is too complex or difficult to understand	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than the group with more than 20 years of experience.	< 2 years: 2.15
		2-5 years: -1.25
		5-10 years: 1.19
		10-20 years: -0.22
		> 20 years: -1.44
Data rights: government requesting proprietary information	The groups of businesses with less than 2 and 2–5 years of experience perceive this to be LESS of a barrier than the group with more than 20 years of experience.	< 2 years: -0.81
		2-5 years: -1.94
		5-10 years: -0.22
		10-20 years: -0.14
		> 20 years: 2.68
Contract documents are too long and complex to understand	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than the groups of 10–20 and more than 20 years of experience.	< 2 years: 2.38
		2-5 years: 0.40
		5-10 years: 0.28
		10-20 years: -1.13
		> 20 years: -1.32

TABLE 2. STATISTICALLY SIGNIFICANT DIFFERENCES BETWEEN LEVELS OF DoD EXPERIENCE, CONTINUED

Barrier	Result	Kruskal-Wallis *z* value
Certified cost data difficult to provide	The groups of businesses with less than 2 and 2–5 years of experience perceive this to be MORE of a barrier than the group with more than 20 years of experience.	< 2 years: 2.31 2-5 years: 1.09 5-10 years: -0.03 10-20 years: 0.00 > 20 years: -2.66
Surety bonds being difficult to obtain	The group of businesses with less than 2 years of experience perceives this to be MORE of a barrier than all four groups with more than 2 years of experience.	< 2 years: 4.01 2-5 years: -0.02 5-10 years: -0.37 10-20 years: -1.62 > 20 years: -1.02
Past-performance rating difficult to obtain	The groups of businesses with less than 2 and 2–5 years of experience perceive this to be MORE of a barrier than the groups of 10–20 and more than 20 years of experience.	< 2 years: 3.90 2-5 years: 2.67 5-10 years: -0.02 10-20 years: -2.85 > 20 years: -2.13
Government accounting standards difficult to meet	All four groups of businesses with less than 20 years of experience perceive this to be MORE of a barrier than the group with more than 20 years of experience.	< 2 years: 2.34 2-5 years: 1.10 5-10 years: 1.93 10-20 years: -0.97 > 20 years: -3.57

Small businesses with less experience found nine of these 10 factors to be more of a barrier than did those businesses with more experience. The pairwise analysis we conducted showed that the point at which differences became statistically significant between groups did vary by factor. Six of the factors were perceived to be more of a challenge by those with less than two years of experience when compared to those with more experience; two factors were perceived to be more of a challenge by all businesses with less than 5 years of experience when compared to those with more. All businesses with less than 20 years of experience found government accounting standards more difficult than those with more than 20 years of defense experience. Four of these factors, which are "extra" requirements unique to defense business, do not come as much of a surprise: certified cost data, government accounting standards, having satisfactory past

performance with the government, and surety bonds. Newer small businesses also view the lack of feedback from proposals and the complexity of documentation as deterrents to pursuing additional defense opportunities. The government requesting data rights or proprietary information was perceived as less of a barrier by small businesses with less than 5 years of DoD experience when compared to those with more than 20 years of experience. This is possibly because they do not have as much proprietary information or they may have more flexibility sharing it to get their foot in the door with government contracts.

Small businesses with extensive defense experience view competition from others as much of a barrier as those just breaking into the defense market.

Equally as important, many factors did not statistically vary by level of experience. Small businesses with extensive defense experience view competition from others as much of a barrier as those just breaking into the defense market. The same holds true for levels of profit and the other factors not mentioned in Table 3.

Size Qualifications

We broke the respondents into four independent groups based on number of employees: 1–10, 11–30, 31–100, and more than 100. We again used Spearman's rho correlation coefficient as a preliminary analysis of the relationship between the factors and number of employees. This time eight of the 26 factors presented correlation coefficients that were statistically significant. Seven of these presented negative correlation coefficients, ranging from -.08 to -.215, indicating a slight trend that the factors become less of an issue as business size increases. However, the trend is fairly weak, as the majority of the factors do not present statistically significant correlation coefficients. The one factor that provided a statistically significant positive trend was "profits too low," with a correlation coefficient of .142.

We again broke out the respondents into four independent groups by number of employees and used the Kruskal-Wallis and Mann-Whitney tests to examine statistically significant differences in the median response. Eight of the 26 factors were statistically significant based on the Mann-Whitney pairwise analysis. These factors and their associated Kruskal-Wallis z values are shown in Table 3.

TABLE 3. STATISTICALLY SIGNIFICANT DIFFERENCES BY COMPANY SIZE

Barrier	Result	Kruskal-Wallis *z* value
Government websites hard to use	The group of businesses with 10 or fewer employees perceives this to be MORE of a barrier than the group with more than 100 employees.	1-10 employees: 1.18
		11-30 employees: 0.20
		31-100 employees: 0.30
		>100 employees: -2.29
Request for proposal/quotation/bid documentation is too complex or difficult to understand	All three groups of businesses with 100 or fewer employees perceive this to be MORE of a barrier than the group with more than 100 employees.	1-10 employees: 0.97
		11-30 employees: 0.33
		31-100 employees: 0.52
		>100 employees: -2.40
Profits too low	The group of businesses with 10 or fewer employees perceives this to be LESS of a barrier than all three groups with more than 10 employees.	1-10 employees: -3.35
		11-30 employees: 1.24
		31-100 employees: 0.93
		>100 employees: 2.18
Contract documents are too long and complex to understand	The groups of businesses with 1–10 and 11–30 employees perceive this to be MORE of a barrier than the group with more than 100 employees.	1-10 employees: 0.75
		11-30 employees: 1.90
		31-100 employees: -0.76
		>100 employees: -2.39
Certified cost data difficult to obtain	The groups of businesses with 1–10 and 11–30 employees perceive this to be MORE of a barrier than the groups with 31-100 and more than 100 employees.	1-10 employees: 2.51
		11-30 employees: 1.34
		31-100 employees: -1.94
		>100 employees: -2.76
Past-performance rating is difficult to obtain	The groups of businesses with 1–10 and 11–30 employees perceive this to be MORE of a barrier than the groups with 31–100 and more than 100 employees.	1-10 employees: 2.61
		11-30 employees: 1.59
		31-100 employees: -2.52
		>100 employees: -2.42

TABLE 3. STATISTICALLY SIGNIFICANT DIFFERENCES BY COMPANY SIZE, CONTINUED		
Barrier	**Result**	**Kruskal-Wallis *z* value**
Government accounting standards difficult to meet	The groups of businesses with 1–10 and 11–30 employees perceive this to be MORE of a barrier than the groups with 31–100 and more than 100 employees.	1-10 employees: 3.74
		11-30 employees: 1.73
		31-100 employees: -3.05
		>100 employees: -3.60
Government regulations are too difficult to understand	The groups of businesses with 1–10 and 11–30 employees perceive this to be MORE of a barrier than the groups with 31–100 and more than 100 employees.	1-10 employees: 1.40
		11-30 employees: 2.81
		31-100 employees: -1.07
		>100 employees: -2.54

Smaller small businesses found seven of these eight factors to be more of a barrier than larger ones. A pairwise analysis again showed that the point at which differences became statistically different did vary some between the factors. Five of the factors, including three of the "extra" requirements specific to government contracts, were perceived to be more of a challenge by businesses with 30 or fewer employees when compared to the larger businesses. Businesses with fewer than 10 employees found government websites and government regulations to be more burdensome when compared to the larger small businesses with more than 100 employees. All businesses with fewer than 100 employees perceived RFP documentation to be complex when compared to businesses with more than 100 employees. Businesses with fewer than 10 employees perceived profits (or lack thereof) as less of a barrier to defense contracts when compared to larger businesses. Again, this possibly indicates that smaller businesses are more willing to sacrifice higher profits to get in the door with government contracts. We found no statistical evidence to support the idea that the larger companies perceive competition from other firms as less of a barrier than those smaller companies with only a handful of employees; the same holds true for the other factors not mentioned in Table 3.

Ideas for a Micro-Business Category

Six factors statistically proved to be more of a barrier by both businesses with less DoD experience and the smaller small businesses: government websites being hard to use, RFP documentation being too complex, contract documents being too complex, certified cost data being hard to provide, past-performance rating being hard to obtain, and government accounting standards being hard to meet. The creation of a defense micro-business category that focuses on reducing these six factors as barriers would likely improve participation from smaller and nontraditional small businesses. Based on the survey results, limiting the size qualifications to approximately 30 employees would be appropriate. Ideas for implementation of this category, provided by survey respondents, are shown below:

- Combine all of the relevant small business DoD websites into a single user-friendly one that shows all steps needed to properly develop and complete proposals. Also consider developing a central website that would help facilitate teaming among complementary small businesses.
- For technical proposals, use white papers as an initial screening process. The initial screening based solely on technical merit will help industry and government focus on technical content and avoid the preparation and review of lengthy proposals, saving both proposers and target agencies time and resources.
- Adopt commercial business practices, particularly commercial-style contracts. Tailor contracts to the greatest extent possible by eliminating clauses that do not apply to the particular contract.
- Ease up on past-performance requirements on these solicitations, particularly when new technology or innovation is important. Evaluate previous commercial experience as an alternative to looking only at past defense contracts on these solicitations.
- Make requirements for government-approved accounting systems less burdensome. Provide accounting software that would allow small businesses to be compliant with cost-reimbursement contracts rather than having them develop or source their own software.

Summary of Findings

We find that most of the factors preventing small businesses from participating in defense contracts identified in previous studies are widely perceived to exist. Of the 26 potential barriers surveyed, most were rather commonly perceived to exist ; 19 of them were rated as a "somewhat important" or "very important" barrier by more than half of

the respondents. Five of the factors provided statistically significant differences in perception between industries: government websites being hard to use, data rights, export control regulations, payment issues, and surety bonds being difficult to obtain. However, 21 of them did not, indicating the majority of the barriers we explored are perceived with no significant difference across industry types.

We did find statistical evidence to support the idea that businesses with less defense business experience perceive defense business to be more challenging than those with extensive defense experience. We also found support for smaller small businesses perceiving defense business to be more challenging than those larger businesses that still qualify for small business contracts. However, this turned out to be the case for only a minority of the factors we explored: nine factors were perceived as more of a barrier by businesses with less defense experience, seven factors were perceived as more of a barrier by the smaller small businesses, and six factors were perceived as more of a barrier by both businesses with less experience and the smaller small businesses. These six were government websites being hard to use, RFP documentation being too complex, contract documents being too complex, certified cost data being hard to provide, past-performance rating being hard to obtain, and government accounting standards being hard to meet.

One way to increase small business participation in defense contracts is to focus reform efforts in areas that small businesses perceive as barriers to defense contracts. The results of this study can be used to concentrate on reducing barriers that will have the largest effect. A concerted attempt to improve communication and response times will likely yield the best results, followed by simplifying the contract proposal process. Additionally, the creation of a micro-business category dedicated to reducing the factors that are more of a barrier to the smaller and newer businesses would likely increase participation in defense contracts by these “nontraditional” small businesses.

Author Note

This article is approved for public release by the Missile Defense Agency, 16-MDA-8666 (13 May 16).

References

Bail, P. Jr. (2010). The demise of the federal government Small Business Program. *Defense Acquisition Research Journal, 17*(1), 77–92.

Blakey, M. (Ed.). (2011, November). *Defense acquisition reform: Moving toward an efficient acquisition system.* Retrieved from http://www.aia-aerospace.org/wp-content/uploads/2016/05/report_acquisition_reform.pdf

Chandler, S. (2014, November 1). *Rethinking competition in defense acquisition.* Retrieved from http://www.lexingtoninstitute.org/wp-content/uploads/2014/11/Rethinking-Competition.pdf

Chaplain, C. (2010, November). *Space aquisitions challenges in commercializing technologies developed under the Small Business Innovation Research program* (GAO-11-21). Retrieved from http://www.gao.gov/assets/320/312130.pdf

Cox, A. G., Moore, N. Y., & Grammich, C. A. (2014). *Identifying and eliminating barriers faced by nontraditional Department of Defense suppliers* (RR-267-OSD). Retrieved from http://www.rand.org/pubs/research_reports/RR267.html

Erwin, S. (2014, January 1). DoD clashes with suppliers over data rights. *National Defense.* Retrieved from http://www.nationaldefensemagazine.org/archive/2014/January/Pages/DoDClashesWithSuppliersOverDataRights.aspx

Freedberg, S. (2014, August 8). *Pentagon struggles to get small-biz tech.* Retrieved from http://breakingdefense.com/2014/08/pentagon-struggles-to-get-small-biz-tech/

Friar, A. (2015, January/February). Swamped by regulations: Perils of an ever-increasing burden. *Defense AT&L, 44*(1), 32–35.

Graves, S. (n.d.). *Fast facts on small business.* Retrieved from http://smallbusiness.house.gov/uploadedfiles/april_recess_small_biz_talking_pts.pdf

House Armed Services Committee. (2012, March 19). *Challenges to doing business with the Department of Defense* (Report). Retrieved from http://wcoeusa.org/sites/default/files/Challenges to Bus with DOD.3.12.pdf

Kendall, F. (2015, April 9). *Implementation directive for Better Buying Power 3.0—Achieving dominant capabilities through technical excellence and innovation* (Memorandum). Retrieved from http://www.acq.osd.mil/fo/docs/betterBuyingPower3.0(9Apr15).pdf

Krieger, J. (2015, January). Why I won't be a prime contractor. *Defense AT&L, 44*(1) 36–39.

Lee, J. (2012, October 9). *DoD carries weight of governmentwide small business goal.* Retrieved from http://federalnewsradio.com/in-depth/2012/10/dod-carries-weight-of-governmentwide-small-business-goal/

Maser, S., & Thompson, F. (2013). Dispelling fear and loathing in government acquisition: A proposal for cultivational governance in DoD source selections. *Journal of Public Procurement, 13*(3), 289–314.

Mielach, D. (2012, May 15). *The big impact of small businesses.* Retrieved from http://www.businessnewsdaily.com/2527-big-impact-small-businesses.html

Mills, S. (2010, March–April). "We don't dance well." *Defense AT&L, 39*(2), 28–33.

Moore, N., Grammich, C., DaVanzo, J., Held, B., Coombs, J., & Mele, J. (2008). *Enhancing small-business opportunities in the DoD.* Retrieved from http://www.rand.org/pubs/technical_reports/TR601-1.html

Neumann, J. (2015, April). *Small business research programs challenges remain in meeting spending and reporting requirements* (GAO-15-358). Retrieved from http://www.gao.gov/assets/670/669650.pdf

Serbu, J. (2015, June 29). *Agencies break record for small business contracting awards.* Retrieved from http://federalnewsradio.com/acquisition/2015/06/government-meets-small-business-goals-second-year-row-beats-previous-records/

Shear, W. (2011, June 1). *Small business contracting action needed by those agencies whose advocates do not report to agency heads as required* (GAO-11-418). Retrieved from http://www.gao.gov/new.items/d11418.pdf

Simmers, E. (2011, September 11). *Ashton Carter talks small business, productivity growth in Orlando.* Retrieved from https://weaponizedculture.wordpress.com/tag/ashton-carter/

Woods, W. (2013, November). *Small business contracting: Updated guidance and reporting needed for consolidated contracts* (GAO-14-36). Retrieved from http://www.gao.gov/assets/660/659254.pdf

Biographies

Mr. Ronnie Schilling works as an engineer with the Missile Defense Agency. He is currently a PhD candidate in Systems Engineering at The George Washington University. Mr. Schilling holds a BS in Mechanical Engineering from The University of Colorado at Colorado Springs and an MS in Aerospace Systems Engineering from The University of Alabama in Huntsville.

(E-mail address: ronnie.schilling@mda.mil)

Dr. Thomas A. Mazzuchi is chair of the Department of Engineering Management and Systems Engineering in the School of Engineering and Applied Science, and professor of Engineering Management and Systems Engineering, at The George Washington University. He earned his BA in Mathematics from Gettysburg College; his MS and DSc in Operations Research from The George Washington University. Dr. Mazzuchi has conducted research for the U.S. Air Force, Army, U.S. Postal Service, and NASA, among others.

(E-mail address: mazzu@gwu.edu)

Dr. Shahram Sarkani is professor of Engineering Management and Systems Engineering, and faculty advisor and academic director of Engineering Management and Systems Engineering off-campus programs at The George Washington University. He designs and administers graduate programs enrolling over 1,000 students across the United States and abroad. Dr. Sarkani earned a BS and MS in Civil Engineering from Louisiana State University, and a PhD in Civil Engineering from Rice University.

(E-mail address: sarkani@gwu.edu)

USING HEURISTICS for Supportability Analysis of Adaptive Weapon Systems in Combat

Samuel H. Amber

The new U.S. Army vision contends that heuristics are practical tools for achieving innovation. Overcoming complex terrain and adaptive hybrid threats in Syria, Iraq, and Afghanistan requires technological innovation. Supportability issues result from modifying deployed weapon systems with new technology for countering these types of threats. Collecting detailed data on deployed weapon systems is constrained in combat zones. A solution for modeling supportability requirements of adaptive weapon systems in a constrained data environment involves heuristics. This modeling effort is achieved by modifying a decision matrix to include heuristics as an alternative field data source.

DOI: *http://dx.doi.org/10.22594/dau.16-743.24.01*
Keywords: *innovation, logistics, decision matrix*

Image designed by Diane Fleischer

Complex terrain and hybrid threats in Syria, Iraq, and Afghanistan are indicative of the United States' future warfare challenges and need for adaptable and innovative weapon systems. In regard to innovation, the new *Army Vision* calls for the ability "to rapidly identify and grapple with complex problems and develop heuristics, or rules of thumb, to adapt and achieve results" (Department of the Army, 2015, p. 8).

Hybrid threats adapt their tactics to counter U.S. conventional military strengths, especially weapon systems such as tanks and infantry carriers. Deployed weapon systems can encounter supportability issues after being modified with new technology that is designed to overcome a hybrid threat. These supportability issues usually include degraded reliability, availability, and maintainability (RAM). However, the hazardous nature of combat zones constrains the collection of detailed sustainment data on modified weapon systems, which can limit effective supportability planning. Thus, decision analysis techniques capable of utilizing alternative and limited data sources are necessary.

A heuristic rule of thumb, educated guess, or trial-and-error result is a useful alternative data source for combat situations where experienced-based, after-action report information is often sufficient for immediate problem solving. Furthermore, a decision matrix correlates and weighs different factors to support decision making, and this analytical technique can be modified to use heuristic-based factors.

Integrating Heuristics into Decision Analysis

This investigation used Figure 1's decision matrix configuration as a starting point. This particular matrix example outlines an analysis for determining the engineering design priorities for a compact disk jewel case. The central portion of the matrix contains relationship strength values between factors located on the left and to the top. The customer's requirements and their importance rating are placed to the left. Data-driven technical characteristics needed to meet the customer's requirements are placed at the top. The absolute and relative weighting of the relationship matrix values are placed at the bottom. The value weighting and rank ordering identify which technical characteristics best fulfill the customer's requirements.

FIGURE 1. DECISION MATRIX EXAMPLE FOR DESIGNING A COMPACT DISK JEWEL CASE

Improvement Direction		↓	↓	**n/a**	↑	**n/a**	**n/a**
	Units	**lbs**	**in**	**n/a**	**ksi$\sqrt{\text{in}}$**	**n/a**	**n/a**
Customer Requirements	Importance weight factor	Force to open	External dimensions	CD positioning feature in case	Toughness of case material	Hinge design	Shape of case
Cost	5		9	3	9	9	
Crack-resistant	5	3	3	3	3	1	3
Ease of stacking	5	3	3			1	
Ease of removing liner notes	5		3				1
Ease of removing CD	4	3	3	9		1	3
Made of recyclable materials	4				9		
Ease of opening case	4	9	3	1		3	3
Scratch resistance	4		1		3		
Hinge stays together	4	3	3			9	3
Waterproof	4	3			3	1	
	Raw Score	102	130	70	120	111	56
	Relative Weight %	17.3	22.1	11.9	20.4	18.8	9.5
	Rank Order	4	1	5	2	3	6

Source: Deiter & Schmidt, 2009, p. 105
Note. The Importance Weight Factor is scaled from 1 to 5, with 1 referring to least important and 5 referring to most important. The central portion of the matrix uses relationship strength values of 9 for strong, 3 for medium, and 1 for weak. Relationship strength values are multiplied by the Importance Weight Factor, and each column is summed to calculate a raw score.

In a limited data situation, a decision matrix with alternative data types, such as the *Army Vision* recommended heuristics, can assist the analysis of supportability issues for an adaptive weapon system baseline. The Army's Stryker vehicle experience in Iraq—documented by Paul Alfieri and Donald McKeon (2008)—presents this type of analysis opportunity.

The Stryker vehicle experienced several baseline changes while deployed initially to Iraq in 2003–2004 to combat an adaptive threat. The deployed vehicle's new technology insertions and extended operations over complex terrain resulted in significant operational suitability issues—such as degraded RAM and increased sustainment costs (Alfieri & McKeon, 2008). Additionally, detailed sustainment data collection was constrained by combat conditions where "recording and reporting data is not a high priority for operational crews" (Alfieri & McKeon, 2008, p. 60).

For this investigation's decision matrix, the customer requirement is structured for maximizing supportability, and it includes the U.S. Army's baseline reliability and availability requirements for the Stryker vehicle. A maintainability requirement was not available from the Stryker program office. Since Stryker data collection was limited in its combat zone, heuristics are placed at the top to represent the environmental and operational use factors affecting supportability.

Table 1's heuristics model the Stryker vehicle's actual supportability expectations in its Iraq operating environment. Specifically, the heuristics are presented as answers to questions about complex terrain and hybrid threats. These questions are based on a widely used U.S. military information format known as the 5W report: Who, What, Where, Why, and When.

TABLE 1. HEURISTICS FOR DEPLOYING STRYKER FORCES AGAINST AN ADAPTIVE HYBRID THREAT

5Ws	Heuristic
Who	Who is deploying to the complex operating environment to fight against an adaptive hybrid threat? Answer: Army Force Generation deployment models are built around the Brigade Combat Team (BCT) force structure. Likewise, this investigation's data set focused on Stryker BCT operations.
What	What kind of Area of Operation (AO) is the Stryker BCT deploying into (and what are its ramifications)? Answer: Higher operational tempo due to an increasingly expanding AO. Note: Work is force applied over a distance, so having more land area to patrol requires more work over time, which directly relates to system usage.
Where	Where will the Stryker BCT conduct operations in terms of terrain? Answer: Primarily operating in complex terrain and urban environments.
Why	Why does the Stryker BCT have to modify its ground combat vehicle system baseline? Answer: Rapid technology insertion to counter an adaptive threat.
When	When will the Stryker BCT conduct operations (i.e., length of deployment and its ramifications)? Answer: Sustainment of adaptive combat vehicle systems during deployments that are far longer than the original Stryker vehicle Objective Requirements Document requirement.

The matrix vertical columns include an Importance (or Weighting) Factor representing the hierarchy of acquisition requirements (i.e., key performance parameters [KPP], key system attributes [KSA], and objective and threshold requirements). An Importance Factor scale of 1, 2, 3, and 4 was assigned for objective, threshold, KSA, and KPP requirements respectively (Wasek, 2005, p. 77). The matrix configuration modified for heuristics and ready for the input of relationship strength values is presented in Figure 2.

FIGURE 2. DECISION MATRIX MODIFIED FOR HEURISTICS AND MAXIMIZING SUPPORTABILITY

		Importance Factor	Higher OPTEMPO due to an increasingly expanding AO	Primarily operating in complex terrain and urban environments	Rapid technology insertion to counter an adaptive threat	Sustainment of adaptive combat vehicle systems during deployments that are longer than the original requirement
Maximize Supportability	Maximize Reliability Reliability Metric = ?					
	Maximize Availability Availability Metric = ?					
	Absolute Weighting					
	Relative Weighting %					
	Rank Order					

Note. The "Who" heuristic is not needed in this figure since the matrix is specific to a weapon system and not the identity of a military unit. OPTEMPO = operations tempo.

Stryker Supportability Analysis

A number value is inputted into each central matrix cell to assess the relationship strength between the U.S. Army's requirements for maximizing supportability and the heuristics associated with the Iraq operating environment. Various value numbering schemes appear in decision matrix literature, and this investigation used the same 9 (strong), 3 (medium), and 1 (weak) scheme outlined in Figure 1 (Deiter & Schmidt, 2009, p. 103). Across each row, the Importance Factor is multiplied by each relationship strength value, and the resulting column values are summed to determine an absolute weight for each heuristic's effect on maximizing supportability. Rank ordering the relative percentile weights identifies the priority of effort for improving supportability.

Adaptive weapon systems usually include Contractor Logistics Support (CLS), and the contract structure normally incentivizes CLS to exceed the KPP Operational Readiness Rate. Though this process generates higher sustainment costs, supplemental funding is usually programmed to support the sustainment needs of contingency deployments. Accordingly, the strongest relationship (9) exists between maximizing availability and the sustainment of adaptive combat vehicle systems during deployments that are longer than the original requirement.

Conversely, weak (1) relationships exist among increased reliability, increased availability, higher operations tempo (OPTEMPO), and complex terrain, since those factors tend to work against one another. Rapid technology insertion has both positive and negative qualities, hence a medium (3) relationship. The positive aspect involves the use of rapid technology insertion to improve the operational effectiveness, suitability, and survivability of weapon system components (e.g., new armor, fire control, power plant, etc.). However, rapid technology insertion tends to decrease a weapon system's overall reliability due to early failures of the newly inserted components (i.e., the first phase of the reliability bathtub curve). Figure 3 presents the completed decision matrix.

FIGURE 3. COMPLETED DECISION MATRIX FOR STRYKER SUPPORTABILITY

		Importance Factor	Higher OPTEMPO due to an increasingly expanding AO	Primarily operating in complex terrain and urban environments	Rapid technology insertion to counter an adaptive threat	Sustainment of adaptive combat vehicle systems during deployments that are longer than the original requirement
Maximize Supportability	Maximize Reliability (KSA) Reliability Metric = 1000 MMBCF	3	1	1	3	1
	Maximize Availability (KPP) Availability Metric ≥ 90% ORR	4	1	1	3	9
	Absolute Weight		7	7	21	39
	Relative Weight (%)		9.46%	9.46%	28.38%	52.70%
	Rank Order		3	3	2	1

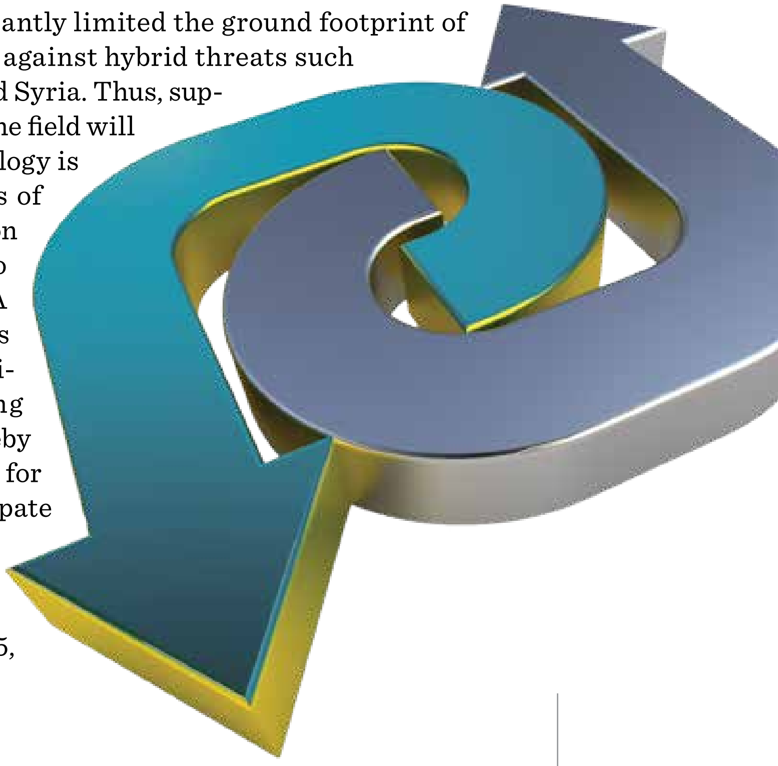
Note. KSA = key system attribute; KPP = key performance parameter; MMBCF = Mean Miles Between Critical Failure; OPTEMPO = operations tempo; ORR = Operational Readiness Rate (Alfieri & McKeon, 2008, p. 55; Department of Defense, 2010, p. 11).

Stryker Supportability Conclusions

Based upon Figure 3's highest relative weight of 52.70%, the recommended priority of effort is to maximize the availability of adaptive combat vehicle systems during deployments that are longer than the original requirement. Likewise, an availability improvement strategy must seek to reduce costs. Since CLS and its higher sustainment costs usually accompany the introduction of new technology into an adaptive baseline, a viable plan for maximizing availability and reducing costs would involve an early design effort for replacing contractor maintenance with soldier mechanics and developing a more robust force structure to manage the Stryker maintenance supply chain better.

After the initial Stryker vehicle deployment to Iraq ended in 2004, the U.S. Army commenced plans to expand significantly Stryker Brigade Combat Team sustainment force structure, including increased soldier mechanic personnel and the activation of forward support companies in each brigade's maneuver battalion (Department of the Army, 2014, pp. 1-10 and C-1; Government Accountability Office, 2006). In hindsight, the use of a decision matrix and heuristics developed from the initial Iraq maintenance reports could have assisted an earlier and similar decision cycle for improving sustainment and supportability of the deployed Stryker vehicle fleet.

President Obama has significantly limited the ground footprint of U.S. military forces deployed against hybrid threats such as the Islamic State of Iraq and Syria. Thus, supportability data collection in the field will be constrained as new technology is provided to smaller numbers of U.S. forces and existing weapon system baselines are adapted to counter hybrid threat tactics. A decision matrix using heuristics presents a method for acquisition supportability planning during current conflicts, thereby serving the Army vision need for "enhancing methods to anticipate future demands on our forces and increased investments in research and development" (Department of the Army, 2015, pp. 8–9).



Acknowledgment

This article is derived from the author's 2010 doctoral research under the advisement of Drs. Julie J. C. H. Ryan, Shahram Sarkani, and Thomas A. Mazzuchi at The George Washington University.

References

Alfieri, P., & McKeon, D. (2008). Stryker suitability challenges in a complex threat environment. *Defense Acquisition Review Journal, 15*(1), 48–63. Retrieved from http://www.dau.mil/pubscats/PubsCats/ARJ47_Alfieri.pdf

Deiter, G. E., & Schmidt, L. C. (2009). *Engineering design* (4th ed.). New York, NY: McGraw-Hill.

Department of Defense. (2010, December). *Selected acquisition report: Stryker.* Retrieved from http://www.dod.mil/pubs/foi/Reading_Room/Selected_Acquisition_Reports/STRYKER-SAR-25_DEC_2010.pdf

Department of the Army. (2014, April). Brigade support battalion*. Army Techniques Publication No. 4-90*. Washington, DC: Author.

Department of the Army. (2015). *The Army vision: Strategic advantage in a complex world.* Retrieved from http://www.army.mil/e2/rv5_downloads/info/references/the_army_vision.pdf

Government Accountability Office. (2006, September). *Defense logistics: Changes to Stryker vehicle maintenance support should identify strategies for addressing implementation challenges* (GAO-06-928R). Retrieved from http://www.gao.gov/assets/100/94423.pdf

Wasek, J. S. (2005). *Mapping system capabilities to requirements and performance metrics for capabilities-based acquisitions: DoD applications* [Doctoral dissertation draft]. The George Washington University.

Biography

Dr. Samuel H. Amber recently completed a 2016 Congressional Fellowship in the U.S. Senate. Dr. Amber was a field artillery officer for 7 years followed by 3 years of program management activities with the U.S. Army's Tactical Exploitation of National Capabilities program. He holds a PhD in Systems Engineering from The George Washington University.

(E-mail address: samuel.amber.army@gmail.com)

The Threat Detection System THAT CRIED WOLF: Reconciling Developers with Operators

Shelley M. Cazares

The Department of Defense and Department of Homeland Security use many threat detection systems, such as air cargo screeners and counter-improvised-explosive-device systems. Threat detection systems that perform well during testing are not always well received by the system operators, however. Some systems may frequently "cry wolf," generating false alarms when true threats are not present. As a result, operators lose faith in the systems—ignoring them or even turning them off and taking the chance that a true threat will not appear. This article reviews statistical concepts to reconcile the performance metrics that summarize a developer's view of a system during testing with the metrics that describe an operator's view of the system during real-world missions. Program managers can still make use of systems that "cry wolf" by arranging them into a tiered system that, overall, exhibits better performance than each individual system alone.

DOI: *http://dx.doi.org/10.22594/dau.16-749.24.01*
Keywords: *probability of detection, probability of false alarm, positive predictive value, negative predictive value, prevalence*

Image designed by Diane Fleischer


320
330
340
350
0
10
20
30
40
50
60
70
80
90

The Department of Defense (DoD) and Department of Homeland Security (DHS) operate many threat detection systems. Examples include counter-mine and counter-improvised-explosive-device (IED) systems and airplane cargo screening systems (Daniels, 2006; L3 Communications Cyterra, 2012; L3 Communications, Security & Detection Systems, 2011, 2013, 2014; Niitek, n.d.; Transportation Security Administration, 2013; U.S. Army, n.d.; Wilson, Gader, Lee, Frigui, & Ho, 2007). All of these systems share a common purpose: to detect threats among clutter.

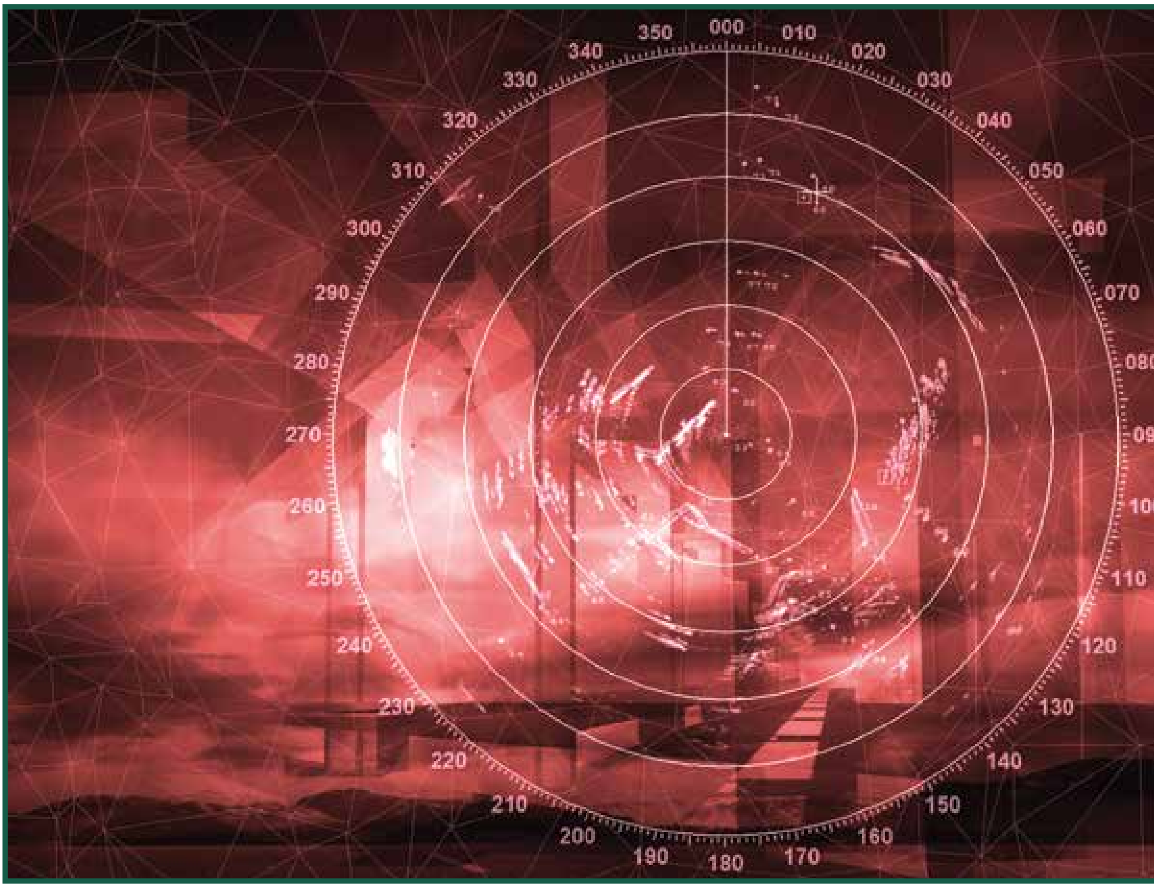
Threat detection systems are often assessed based on their Probability of Detection (P_d) and Probability of False Alarm (P_{fa}). P_d describes the fraction of true threats for which the system correctly declares an alarm. Conversely, P_{fa} describes the fraction of true clutter (true non-threats) for which the system *in*correctly declares an alarm—a false alarm. A perfect system will exhibit a P_d of 1 and a P_{fa} of 0. P_d and P_{fa} are summarized in Table 1 and discussed in Urkowitz (1967).

TABLE 1. DEFINITIONS OF COMMON METRICS USED TO ASSESS PERFORMANCE OF THREAT DETECTION SYSTEMS

Metric	Definition	Perspective
Probability of Detection (P_d)	The fraction of all items containing a true threat for which the system correctly declared an alarm	Developer
Probability of False Alarm (P_{fa})	The fraction of all items *not* containing a true threat for which the system *in*correctly declared an alarm	Developer
Positive Predictive Value (PPV)	The fraction of all items causing an alarm that did end up containing a true threat	Operator
Negative Predictive Value (NPV)	The fraction of all items *not* causing an alarm that did end up *not* containing a true threat	Operator
Prevalence (Prev)	The fraction of items that contained a true threat (regardless of whether the system declared an alarm)	—
False Alarm Rate (FAR)	The number of false alarms per unit time, area, or distance	—

Threat detection systems with good P_d and P_{fa} performance metrics are not always well received by the system's operators, however. Some systems may frequently "cry wolf," generating false alarms when true threats are not present. As a result, operators may lose faith in the systems, delaying their response to alarms (Getty, Swets, Pickett, & Gonthier, 1995) or ignoring

them altogether (Bliss, Gilson, & Deaton, 1995), potentially leading to disastrous consequences. This issue has arisen in military, national security, and civilian scenarios.



The New York Times described a 1987 military incident involving the threat detection system installed on a $300 million high-tech warship to track radar signals in the waters and airspace off Bahrain. Unfortunately, "somebody had turned off the audible alarm because its frequent beeps bothered him" (Cushman, 1987, p. 1). The radar operator was looking away when the system flashed a sign alerting the presence of an incoming Iraqi jet. The attack killed 37 sailors.

That same year, *The New York Times* reported a similar civilian incident in the United States. An Amtrak train collided near Baltimore, Maryland, killing 15 people and injuring 176. Investigators found that an alarm whistle

in the locomotive cab had been "substantially disabled by wrapping it with tape" and "train crew members sometimes muffle the warning whistle because the sound is annoying" (Stuart, 1987, p. 1).

Such incidents continued to occur two decades later. In 2006, *The Los Angeles Times* described an incident in which a radar air traffic control system at Los Angeles International Airport (LAX) issued a false alarm, prompting the human controllers to "turn off the equipment's aural alert" (Oldham, 2006, p. 2). Two days later, a turboprop plane taking off from the airport narrowly missed a regional jet, the "closest call on the ground at LAX" in 2 years (Oldham, 2006, p. 2). This incident had homeland security implications, since DHS and the Department of Transportation are co-sector-specific agencies for the Transportation Systems Sector, which governs air traffic control (DHS, 2016).

The disabling of threat detection systems due to false alarms is troubling. This behavior often arises from an inappropriate choice of metrics used to assess the system's performance during testing. While P_d and P_{fa} encapsulate the *developer's* perspective of the system's performance, these metrics do not encapsulate the *operator's* perspective. The operator's view can be better summarized with other metrics, namely Positive Predictive Value (PPV) and Negative Predictive Value (NPV). PPV describes the fraction of all alarms that correctly turn out to be true threats—a measure of how often the system does not "cry wolf." Similarly, NPV describes the fraction of all *lack* of alarms that correctly turn out to be true clutter. From the operator's perspective, a perfect system will have PPV and NPV values equal to 1. PPV and NPV are summarized in Table 1 and discussed in Altman and Bland (1994b).

Interestingly enough, the very same threat detection system that satisfies the developer's desire to detect as much truth as possible can also disappoint the operator by generating false alarms, or "crying wolf," too often (Scheaffer & McClave, 1995). A system

can exhibit excellent P_d and P_{fa} values while also exhibiting a poor PPV value. Unfortunately, low PPV values naturally occur when the Prevalence (Prev) of true threat among true clutter is extremely low (Parasuraman, 1997; Scheaffer & McClave, 1995), as is often the case in defense and homeland security scenarios. As summarized in Table 1, Prev is a measure of how widespread or common the true threat is. A Prev of 1 indicates a true threat is always present, while a Prev of 0 indicates a true threat is never present. As will be shown, a low Prev can lead to a discrepancy in how developers and operators view the performance of threat detection systems in the DoD and DHS.

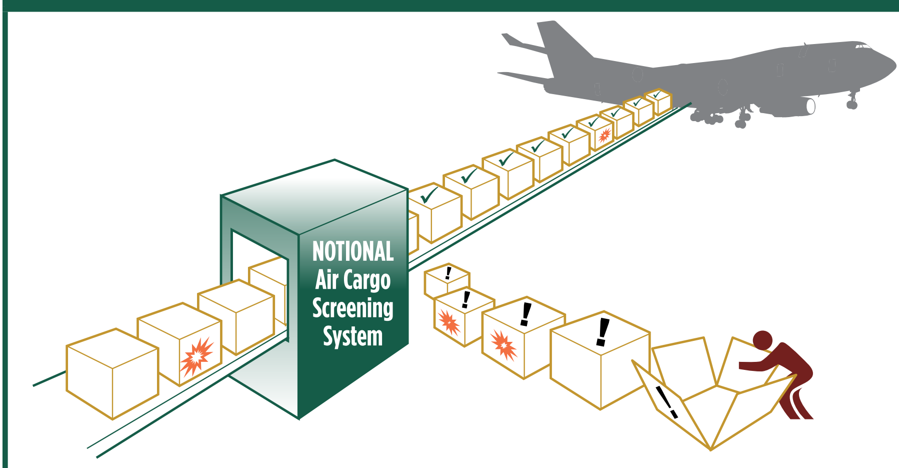
In this article, the author reconciles the performance metrics used to quantify the developer's versus operator's views of threat detection systems. Although these concepts are already well known within the statistics and human factors communities, they are not often immediately understood in the DoD and DHS science and technology (S&T) acquisition communities. This review is intended for program managers (PM) of threat detection systems in the DoD and DHS. This article demonstrates how to calculate P_d, P_{fa}, PPV, and NPV using a notional air cargo screening system as an example. Then it illustrates how a PM can still make use of a system that frequently "cries wolf" by incorporating it into a tiered system that, overall, exhibits better performance than each individual system alone. Finally, the author cautions that P_{fa} and NPV can be calculated only for threat *classification* systems, rather than genuine threat *detection* systems. False Alarm Rate is often calculated in place of P_{fa}.

Testing a Threat Detection System

A notional air cargo screening system illustrates the discussion of performance metrics for threat detection systems. As illustrated by Figure 1, the purpose of this notional system is to detect explosive threats packed inside items that are about to be loaded into the cargo hold of an airplane. To determine how well this system meets capability requirements, its performance must be quantified. A large number of items is input into the system, and each item's ground truth (whether the item contained a true threat) is compared to the system's output (whether the system declared an alarm). The items are representative of the items that the system would likely encounter in an operational setting. At the end of the test, the True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) items are counted. Figure 2 tallies these counts in a 2 × 2 confusion matrix:

- A TP is an item that contained a true threat, and for which the system correctly declared an alarm.
- An FP is an item that did *not* contain a true threat, but for which the system *in*correctly declared an alarm—a false alarm (a Type I error).
- An FN is an item that contained a true threat, but for which the system *in*correctly did *not* declare an alarm (a Type II error).
- A TN is an item that did *not* contain a true threat, and for which the system correctly did *not* declare an alarm.

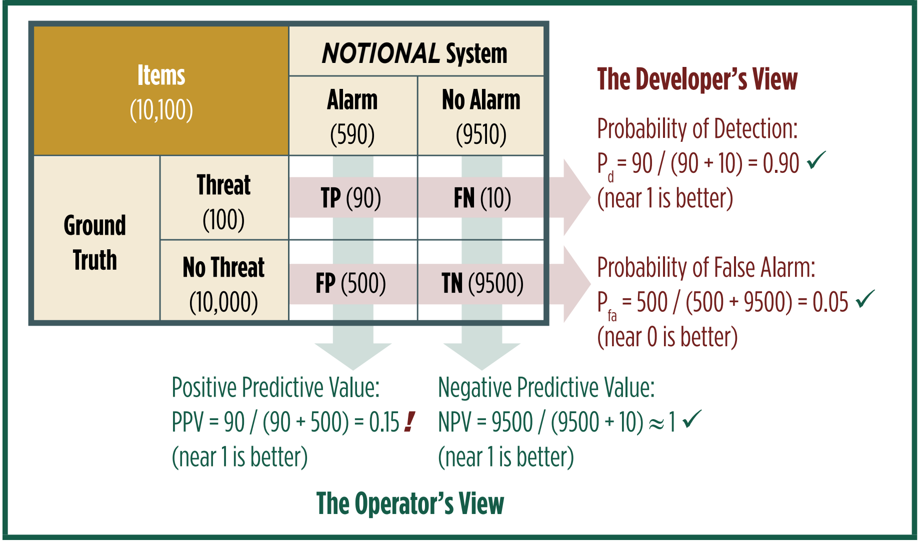
FIGURE 1. NOTIONAL AIR CARGO SCREENING SYSTEM



Note. A set of predefined, discrete items (small brown boxes) are presented to the system one at a time. Some items contain a true threat (orange star) among clutter, while other items contain clutter only (no orange star). For each item, the system declares either one or zero alarms. All items for which the system declares an alarm (black exclamation point) are further examined manually by trained personnel (red figure). In contrast, all items for which the system does not declare an alarm (green checkmark) are left unexamined and loaded directly onto the airplane.

As shown in Figure 2, a total of 10,100 items passed through the notional air cargo screening system. One hundred items contained a true threat while 10,000 items did not. The system declared an alarm for 590 items and did not declare an alarm for 9,510 items. Comparing the items' ground truth to the system's alarms (or lack thereof), there were 90 TPs, 10 FNs, 500 FPs, and 9,500 TNs.

FIGURE 2. 2 X 2 CONFUSION MATRIX OF NOTIONAL AIR CARGO SCREENING SYSTEM



Note. The matrix tabulates the number of TP, FN, FP, and TN items processed by the system. P_d and P_{fa} summarize the developer's view of the system's performance while PPV and NPV summarize the operator's view. In this notional example, the low PPV of 0.15 indicates a poor operator experience (the system often generates false alarms and "cries wolf," since only 15% of alarms turn out to be true threats) even though the good P_d and P_{fa} are well received by developers.

The Developer's View: P_d and P_{fa}

A PM must consider how much of the truth the threat detection system is able to identify. This can be done by considering the following questions: Of those items that contain a true threat, for what fraction does the system correctly declare an alarm? And of those items that do *not* contain a true threat, for what fraction does the system *in*correctly declare an alarm—a false alarm? These questions often guide developers during the research and development phase of a threat detection system.

P_d and P_{fa} can be easily calculated from the 2 × 2 confusion matrix to answer these questions. From a developer's perspective, this notional air cargo screening system exhibits good[1] performance:

$$P_d = \frac{TP}{TP + FN} = \frac{90}{90 + 10} = 0.90 \text{ (compared to 1 for a perfect system)} \qquad (1)$$

$$P_{fa} = \frac{FP}{FP + TN} = \frac{500}{500 + 9{,}500} = 0.05 \text{ (compared to 0 for a perfect system)} \qquad (2)$$

Equation 1 shows that, of all items that contained a true threat (TP + FN = 90 + 10 = 100), a large subset (TP = 90) correctly caused an alarm. These counts resulted in $P_d = 0.90$, close to the value of 1 that would be exhibited by a perfect system.[2] Based on this P_d value, the PM can conclude that 90% of items that contained a true threat correctly caused an alarm, which may (or may not) be considered acceptable within the capability requirements for the system. Furthermore, Equation 2 shows that, of all items that did *not* contain a true threat (FP + TN = 500 + 9,500 = 10,000), only a small subset (FP = 500) caused a false alarm. These counts led to $P_{fa} = 0.05$, close to the value of 0 that would be exhibited by a perfect system.[3] In other words, only 5% of items that did *not* contain a true threat caused a false alarm.

The Operator's View: PPV and NPV

The PM must also anticipate the operator's view of the threat detection system. One way to do this is to answer the following questions: Of those items that caused an alarm, what fraction turned out to contain a true threat (i.e., what fraction of alarms turned out *not* to be false)? And of those items that did *not* cause an alarm, what fraction turned out *not* to contain a true threat? On the surface, these questions seem similar to those posed previously for P_d and P_{fa}. Upon closer examination, however, they are quite different. While P_d and P_{fa} summarize how much of the truth causes an alarm, PPV and NPV summarize how many alarms turn out to be true.

PPV and NPV can also be easily calculated from the 2 × 2 confusion matrix. From an operator's perspective, the notional air cargo screening system exhibits a conflicting performance:

$$\text{NPV} = \frac{\text{TN}}{\text{TN}+\text{FN}} = \frac{9{,}500}{9{,}500+10} \approx 1 \text{ (compared to 1 for a perfect system)} \quad (3)$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP}+\text{FP}} = \frac{90}{90+500} = 0.15 \text{ (compared to 1 for a perfect system)} \quad (4)$$

Equation 3 shows that, of all items that did *not* cause an alarm (TN + FN = 9,500 + 10 = 9,510), a very large subset (TN = 9,500) correctly turned out to *not* contain a true threat. These counts resulted in NPV ≈ 1, approximately equal to the 1 value that would be exhibited by a perfect system.[4] In the absence of an alarm, the operator could rest assured that a threat was highly unlikely. However, Equation 4 shows that, of all items that did indeed cause an alarm (TP + FP = 90 + 500 = 590), only a small subset (TP = 90) turned out to contain a true threat (i.e., were not false alarms). These counts unfortunately led to PPV = 0.15, much lower than the 1 value that would be

exhibited by a perfect system.[5] When an alarm was declared, the operator could not trust that a threat was present, since the system generated false alarms so often.

Reconciling Developers with Operators: P_d and P_{fa} Versus PPV and NPV

The discrepancy between PPV and NPV versus P_d and P_{fa} reflects the discrepancy between the operator's and developer's views of the threat detection system. Developers are often primarily interested in how much of the truth correctly cause alarms—concepts quantified by P_d and P_{fa}. In contrast, operators are often primarily concerned with how many alarms turn out to be true—concepts quantified by PPV and NPV. As shown in Figure 2, the very same system that exhibits good values for P_d, P_{fa}, and NPV can also exhibit poor values for PPV.

Poor PPV values should not be unexpected for threat detection systems in the DoD and DHS. Such performance is often merely a reflection of the low Prev of true threats among true clutter that is not uncommon in defense and homeland security scenarios.[6] Prev describes the fraction of all items that contain a true threat, including those that did and did not cause an alarm. In the case of the notional air cargo screening system, Prev is very low:

$$\text{Prev} = \frac{\text{TP} + \text{FN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} = \frac{90 + 10}{90 + 10 + 500 + 9{,}500} = 0.01 \qquad (5)$$

Equation 5 shows that, of all items (TP + FN + FP + TN = 90 + 10 + 500 + 9,500 = 10,100), only a very small subset (TP + FN = 90 + 10 = 100) contained a true threat, leading to Prev = 0.01. When true threats are rare, most alarms turn out to be false, even for an otherwise strong threat detection system, leading to a low value for PPV (Altman & Bland, 1994b). In fact, to achieve a high value of PPV when Prev is extremely low, a threat detection system must exhibit so few FPs (false alarms) as to make P_{fa} approximately zero.

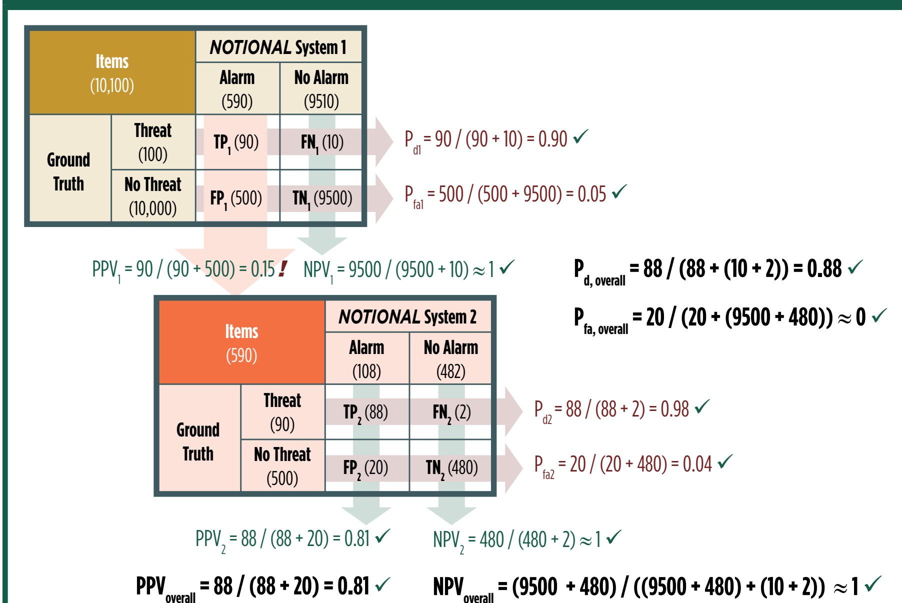
Recognizing this phenomenon, PMs should not necessarily dismiss a threat detection system simply because it exhibits a poor PPV, provided that it also exhibits an excellent P_d and P_{fa}. Instead, PMs can estimate Prev to help determine how to guide such a system through development. Prev does *not* depend on the threat detection system and can, in fact, be calculated in the absence of the system. Knowledge of ground truth (which items contain a true threat) is all that is needed to calculate Prev (Scheaffer & McClave, 1995).

Of course, ground truth is not known *a priori* in an operational setting. However, it may be possible for PMs to use historical data or intelligence tips to roughly estimate whether Prev is likely to be particularly low in operation. The threat detection system can be thought of as one system in a system of systems, where other relevant systems are based on record keeping (to provide historical estimates of Prev) or intelligence (to provide tips to help estimate Prev). These estimates of Prev can vary over time and location. A Prev that is estimated to be very low can cue the PM to anticipate discrepancies in P_d and P_{fa} versus PPV, forecasting the inevitable discrepancy between the developer's versus operator's views early in the system's development, while there are still time and opportunity to make adjustments. At that point, the PM can identify a concept of operations (CONOPS) in which the system can still provide value to the operator for an assigned mission. A tiered system may provide one such opportunity.

A Tiered System for Threat Detection

Tiered systems consist of multiple systems used in series. The first system cues the use of the second system and so on. Tiered systems provide PMs the opportunity to leverage multiple threat detection systems that, individually, do not satisfy both developers and operators simultaneously. Figure 3 shows two 2 × 2 confusion matrices that represent a notional tiered system that makes use of two individual threat detection systems. The first system (top) is relatively simple (and inexpensive) while the second system (bottom) is more complex (and expensive). Other tiered systems can consist of three or more individual systems.

FIGURE 3. NOTIONAL TIERED SYSTEM FOR AIR CARGO SCREENING



Note. The top 2 × 2 confusion matrix represents the same notional system described in Figures 1 and 2. While this system exhibits good P_d, P_{fa}, and NPV values, its PPV value is poor. Nevertheless, this system can be used to cue a second system to further analyze the questionable items. The bottom matrix represents the second notional system. This system exhibits a good P_d, P_{fa}, and NPV, along with a much better PPV. The second system's better PPV reflects the higher Prev of true threat encountered by the second system, due to the fact that the first system had already successfully screened out most items that did not contain a true threat. Overall, the tiered system exhibits a more nearly optimal balance of P_d, P_{fa}, NPV, and PPV than either of the two systems alone.

The first system is the notional air cargo screening system discussed previously. Although this system exhibits good performance from the developer's perspective (high P_d and low P_{fa}), it exhibits conflicting performance from the operator's perspective (high NPV but low PPV). Rather than using this system to classify items as either "Alarm (Threat)" or "No Alarm (No Threat)," the operator can use this system to *screen* items as either "Cue Second System (*Maybe* Threat)" or "Do Not Cue Second System (No Threat)." Of the 10,100 items that passed through the first system, 590 were classified as "Cue Second System (*Maybe* Threat)" while 9,510 were classified as "No Alarm (No Threat)." The first system's extremely high NPV (approximately equal to 1) means that the operator can rest assured that the lack of a cue correctly indicates the very low likelihood of a true threat. Therefore, any item that fails to elicit a cue can be loaded onto the airplane, bypassing the second system and avoiding its unnecessary complexities and expense.[7] In contrast, the first system's low PPV indicates that the operator cannot trust that a cue indicates a true threat. Any item that elicits a cue from the first system may or may not contain a true threat and must therefore pass through the second system for further analysis.

Only 590 items elicited a cue from the first system and passed through the second system. Ninety items contained a true threat, while 500 items did not. The second system declared an alarm for 108 items and did not declare an alarm for 482 items. Comparing the items' ground truth to the second system's alarms (or lack thereof), there were 88 TPs, 2 FNs, 20 FPs, and 480 TNs. On its own, the second system exhibits a higher P_d and lower P_{fa} than the first system, due to its increased complexity (and expense). In addition, its PPV value is much higher. The second system's higher PPV may be due to its higher complexity or may simply be due to the fact that the second system encounters a higher Prev of true threat among true clutter than the first system. By the very nature in which the tiered system was assembled, the first system's very high NPV indicates its strong ability to screen out most items that do *not* contain a true threat, leaving only those questionable items for the second system to process. Since the second system encounters

only those items that are questionable, it encounters a much higher Prev and therefore has the opportunity to exhibit higher PPV values. The second system simply has less relative opportunity to generate false alarms.

> ***The utility of the tiered system must be considered in light of its cost.***

The utility of the tiered system must be considered in light of its cost. In some cases, the PM may decide that the first system is not needed, since the second, more complex, system can exhibit the desired P_d, P_{fa}, PPV, and NPV values on its own. In that case, the PM may choose to abandon the first system and pursue a single-tier approach based solely on the second system. In other cases, the added complexity of the second system may require a large increase in resources for its operation and maintenance. In these cases, the PM may opt for the tiered approach, in which use of the first system reduces the number of items that must be processed by the second system, reducing the additional resources needed to operate and maintain the second system to a level that may balance out the increase in resources needed to operate and maintain a tiered approach.

To consider the utility of the tiered system, its performance as a whole must be assessed, in addition to the performance of each of the two individual systems that compose it. As with any individual system, P_d, P_{fa}, PPV, and NPV can be calculated for the tiered system overall. These calculations must be based on *all* items encountered by the tiered system as a whole, taking care *not* to double count those TP_1 and FP_1 items from the first tier that pass to the second:

$$P_d = \frac{TP_2}{TP_2 + (FN_1 + FN_2)} = \frac{88}{88 + (10 + 2)} = 0.88 \text{ (compared to 1 for a perfect system)} \quad (6)$$

$$P_{fa} = \frac{FP_2}{FP_2 + (TN_1 + TN_2)} = \frac{20}{20 + (9{,}500 + 480)} \approx 0 \text{ (compared to 0 for a perfect system)} \quad (7)$$

$$NPV = \frac{(TN_1 + TN_2)}{(TN_1 + TN_2) + (FN_1 + FN_2)} = \frac{(9{,}500 + 480)}{(9{,}500 + 480) + (10 + 2)} \approx 1 \text{ (compared to 1 for a perfect system)} \quad (8)$$

$$PPV = \frac{TP_2}{TP_2 + FP_2} = \frac{88}{88 + 20} = 0.81 \text{ (compared to 1 for a perfect system)} \quad (9)$$

Overall, the tiered system exhibits good[8] performance from the developer's perspective. Equation 6 shows that, of all items that contained a true threat (TP_2 + (FN_1 + FN_2) = 88 + (10 + 2) = 100), a large subset (TP_2 = 88) correctly caused an alarm, resulting in an overall value of P_d = 0.88. The PM can conclude that 88% of items containing a true threat correctly led to a final alarm from the tiered system as a whole. Although this overall P_d is slightly lower than the P_d of each of the two individual systems, the overall value is still close to the value of 1 for a perfect system[9] and may (or may not) be considered acceptable within the capability requirements for the envisioned CONOPS. Similarly, Equation 7 shows that, of all items that did *not* contain a true threat (FP_2 + (TN_1 + TN_2) = 20 + (9,500 + 480) = 10,000), only a very small subset (FP_2 = 20) *in*correctly caused an alarm, leading to an overall value of $P_{fa} \approx 0$. Approximately 0% of items *not* containing a true threat caused a false alarm.

The tiered system also exhibits good[10] overall performance from the operator's perspective. Equation 8 shows that, of all items that did *not* cause an alarm ((TN_1 + TN_2) + (FN_1 + FN_2) = (9,500 + 480) + (10 + 2) = 9,992), a very large subset ((TN_1 + TN_2) = (9,500 + 480) = 9,980) correctly turned out *not* to contain a true threat, resulting in an overall value of NPV ≈ 1. The operator could rest assured that a threat was highly unlikely in the absence of a final alarm. More interesting, though, is the overall PPV value. Equation 9 shows that, of all items that did indeed cause a final alarm ((TP_2 + FP_2) = (88 + 20) = 108), a large subset (TP_2 = 88) correctly turned out to contain a true threat—these alarms were *not* false. These counts resulted in an overall value of PPV = 0.81, much closer to the 1 value of a perfect system and much higher than the PPV of the first system alone.[11] When a final alarm was declared, the operator could trust that a true threat was indeed present since, overall, the tiered system did not "cry wolf" very often.

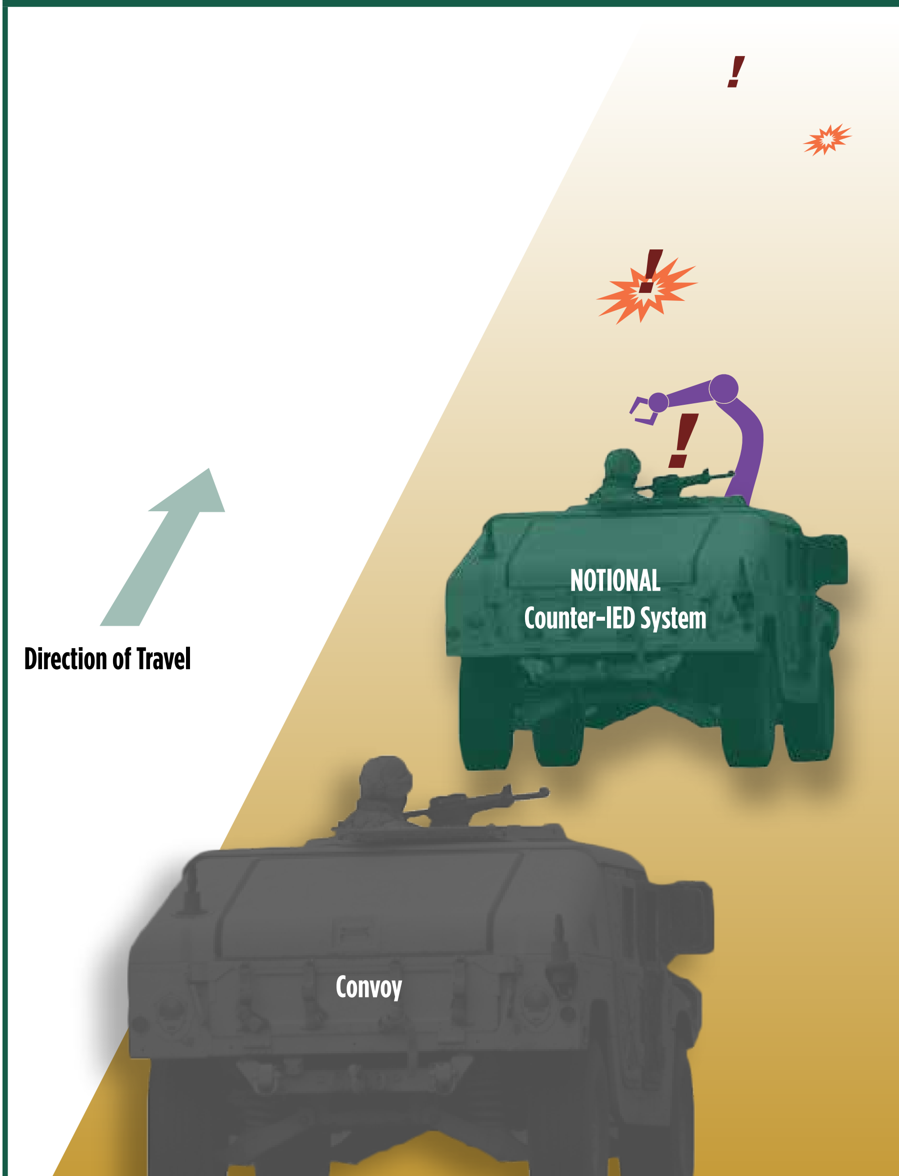
Of course, the PM must compare the overall performance of the tiered system to capability requirements in order to assess its appropriateness for the envisioned mission (DoD, 2015; DHS, 2008). The overall values of P_d = 0.88, $P_{fa} \approx 0$, NPV ≈ 1, and PPV = 0.81 may or may not be adequate once these values are compared to such requirements. Statistical tests can determine whether the overall values of the tiered system are significantly less than required (Fleiss, Levin, & Paik, 2013). Requirements should be set for all four metrics based on the envisioned mission. Setting metrics for only P_d and P_{fa} effectively ignores the operator's view, while setting metrics for only PPV and NPV effectively ignores the developer's view.[12] One may argue that only the operator's view (PPV and NPV) must be quantified as capability requirements. However, there is value in also retaining the developer's view

(P_d and P_{fa}), since P_d and P_{fa} can be useful when comparing and contrasting the utility of rival systems with similar PPV and NPV values in a particular mission. Setting the appropriate requirements for a particular mission is a complex process and is beyond the scope of this article.

Threat Detection Versus Threat Classification

Unfortunately, all four performance metrics cannot be calculated for some threat detection systems. In particular, it may be impossible to calculate P_{fa} and NPV. This is due to the fact that the term "threat detection system" can be a misnomer, because it is often used to refer to threat detection *and* threat classification systems. Threat classification systems are those that are presented with a set of predefined, discrete items. The system's task is to classify each item as either "Alarm (Threat)" or "No Alarm (No Threat)." The notional air cargo screening system discussed in this article is actually an example of a threat *classification* system, despite the fact that the author has colloquially referred to it as a threat *detection* system throughout the first half of this article. In contrast, genuine threat detection systems are those that are *not* presented with a set of predefined, discrete items. The system's task is *first to detect* the discrete items from a continuous stream of data *and then to classify* each detected item as either "Alarm (Threat)" or "No Alarm (No Threat)." An example of a genuine threat detection system is the notional counter-IED system illustrated in Figure 4.

FIGURE 4. NOTIONAL COUNTER-IED SYSTEM



Note. Several items are buried in a road often traveled by a U.S. convoy. Some items are IEDs (orange stars), while others are simply rocks, trash, or other discarded items. The system continuously collects data while traveling over the road ahead of the convoy and declares one alarm (red exclamation point) for each location at which it detects a buried IED. All locations for which the system declares an alarm are further examined with robotic systems (purple arm) operated remotely by trained personnel. In contrast, all parts of the road for which the system does not declare an alarm are left unexamined and are directly traveled over by the convoy.

This issue is more than semantics. Proper labeling of a system's task helps to ensure that the appropriate performance metrics are used to assess the system. In particular, while P_{fa} and NPV can be used to describe threat *classification* systems, they cannot be used to describe genuine threat *detection* systems. For example, Equation 2 showed that P_{fa} depends on FP and TN counts. While an FP is a true clutter item that *in*correctly caused an alarm, a TN is a true clutter item that correctly did *not* cause an alarm. FPs and TNs can be counted for threat *classification* systems and used to calculate P_{fa}, as described earlier for the notional air cargo screening system.

> ***This issue is more than semantics. Proper labeling of a system's task helps to ensure that the appropriate performance metrics are used to assess the system.***

This story changes for genuine threat *detection* systems, however. While FPs can be counted for genuine threat detection systems, TNs cannot. Therefore, while P_d and PPV can be calculated for genuine threat detection systems, P_{fa} and NPV cannot, since they are based on the TN count. For the notional counter-IED system, an FP is a location on the road for which a true IED is *not* buried but for which the system *in*correctly declares an alarm. Unfortunately, a converse definition for TNs does not make sense: How should one count the number of locations on the road for which a true IED is *not* buried and for which the system correctly does *not* declare an alarm? That is, how often should the system get credit for declaring nothing when nothing was truly there? To answer these TN-related questions, it may be possible to divide the road into sections and count the number of sections for which a true IED is *not* buried and for which the system correctly does *not* declare an alarm. However, such a method simply converts the counter-IED *detection* problem into a counter-IED *classification* problem, in which discrete items (sections of road) are predefined and the system's task is merely to classify each item (each section of road) as either "Alarm (IED)" or "No Alarm (No IED)." This method imposes an artificial definition on the item (section of road) under classification: How long should each section of road be? Ten meters long? One meter long? One centimeter long? Such definitions can be artificial, which simply highlights the fact that the concept of a TN does not exist for genuine threat detection systems.

Therefore, PMs often rely on an additional performance metric for genuine threat detection systems—the False Alarm Rate (FAR). FAR can often be confused with both P_{fa} and PPV. In fact, documents within the defense and homeland security communities can erroneously use two or even all three of these terms interchangeably. In this article, however, FAR refers to the number of FPs processed per unit time interval, or unit geographical area, or distance (depending on which metric—time, area, or distance—is more salient to the envisioned CONOPS):

$$FAR = \frac{FP}{total\,time} \tag{10a}$$

or

$$FAR = \frac{FP}{total\,area} \tag{10b}$$

or

$$FAR = \frac{FP}{total\,distance} \tag{10c}$$

For example, Equation 10c shows that one could count the number of FPs processed *per meter* as the notional counter-IED system travels down the road. In that case, FAR would have units of m^{-1}. In contrast, P_d, P_{fa}, PPV, and NPV are dimensionless quantities. FAR can be a useful performance metric in situations for which P_{fa} cannot be calculated (such as for genuine threat detection systems) or for which it is prohibitively expensive to conduct a test to fill out the full 2 × 2 confusion matrix needed to calculate P_{fa}.

Conclusions

Several metrics can be used to assess the performance of a threat detection system. P_d and P_{fa} summarize the developer's view of the system, quantifying how much of the truth causes alarms. In contrast, PPV and NPV summarize the operator's perspective, quantifying how many alarms turn out to be true. The same system can exhibit good values for P_d and P_{fa} during testing but poor PPV values during operational use. PMs can still make use of the system as part of a tiered system that, overall, exhibits better performance than each individual system alone.

References

Altman, D. G., & Bland, J. M. (1994a). Diagnostic tests 1: Sensitivity and specificity. *British Medical Journal, 308*(6943), 1552. doi:10.1136/bmj.308.6943.1552

Altman, D. G., & Bland, J. M. (1994b). Diagnostic tests 2: Predictive values. *British Medical Journal, 309*(6947), 102. doi:10.1136/bmj.309.6947.102

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behavior in response to alarms of varying reliability. *Ergonomics, 38*(11), 2300–2312. doi:10.1080/00140139508925269

Cushman, J. H. (1987, June 21). Making arms fighting men can use. *The New York Times.* Retrieved from http://www.nytimes.com/1987/06/21/business/making-arms-fighting-men-can-use.html

Daniels, D. J. (2006). A review of GPR for landmine detection. *Sensing and Imaging: An International Journal, 7*(3), 90–123. Retrieved from http://link.springer.com/article/10.1007%2Fs11220-006-0024-5

Department of Defense. (2015, January 7). *Operation of the defense acquisition system* (Department of Defense Instruction [DoDI] 5000.02). Washington, DC: Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics. Retrieved from http://bbp.dau.mil/docs/500002p.pdf

Department of Homeland Security. (2008, November 7). *Acquisition instruction/guidebook* (DHS Publication No. 102-01-001, Interim, Version 1.9). Retrieved from http://www.it-aac.org/images/Acquisition_Instruction_102-01-001_-_Interim_v1.9_dtd_11-07-08.pdf

Department of Homeland Security. (2016, March 30). *Transportation systems sector.* Retrieved from https://www.dhs.gov/transportation-systems-sector

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: John Wiley.

Getty, D. J., Swets, J. A., Pickett, R. M., & Gonthier, D. (1995). System operator response to warnings of danger: A laboratory investigation of the effects of the predictive value of a warning on human response time. *Journal of Experimental Psychology: Applied, 1*(1), 19–33. doi:10.1037/1076-898X.1.1.19

L3 Communications Cyterra. (2012). *AN/PSS-14 mine detection.* Orlando, FL: Author. Retrieved from http://cyterra.com/products/anpss14.htm

L3 Communications, Security & Detection Systems. (2011). *Fact sheet: Examiner 3DX explosives detection system.* Woburn, MA: Author. Retrieved from http://www.sds.l-3com.com/forms/English-pdfdownload.htm?DownloadFile=PDF-13

L3 Communications, Security & Detection Systems. (2013). *Fact sheet: Air cargo screening solutions: Regulator-qualified detection systems.* Woburn, MA: Author. Retrieved from http://www.sds.l-3com.com/forms/English-pdfdownload.htm?DownloadFile=PDF-50

L3 Communications, Security & Detection Systems. (2014). *Fact sheet: Explosives detection systems: Regulator-approved checked baggage solutions.* Woburn, MA: Author. Retrieved from http://www.sds.l-3com.com/forms/English-pdfdownload.htm?DownloadFile=PDF-17

Niitek. (n.d.). *Counter IED | Husky Mounted Detection System (HMDS)*. Sterling, VA: Author. Retrieved from http://www.niitek.com/~/media/Files/N/Niitek/documents/hmds.pdf

Oldham, J. (2006, October 3). Outages highlight internal FAA rift. *The Los Angeles Times.* Retrieved from http://articles.latimes.com/2006/oct/03/local/me-faa3

Parasuraman, R. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. doi:10.1518/001872097778543886

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies, 2*(1), 37–63.

Scheaffer, R. L., & McClave, J. T. (1995). Conditional probability and independence: Narrowing the table. In *Probability and statistics for engineers* (4th ed., pp. 85–92). Belmont, CA: Duxbury Press.

Stuart, R. (1987, January 8). U.S. cites Amtrak for not conducting drug tests. *The New York Times.* Retrieved from http://www.nytimes.com/1987/01/08/us/us-cites-amtrak-for-not-conducting-drug-tests.html

Transportation Security Administration. (2013). *TSA air cargo screening technology list (ACSTL)* (Version 8.4 as of 01/31/2013). Washington, DC: Author. Retrieved from http://www.cargosecurity.nl/wp-content/uploads/2013/04/nonssi_acstl_8_4_jan312013_compliant.pdf

Wilson, J. N., Gader, P., Lee, W. H., Frigui, H., and Ho, K. C. (2007). A large-scale systematic evaluation of algorithms using ground-penetrating radar for landmine detection and discrimination. *IEEE Transactions on Geoscience and Remote Sensing, 45*(8), 2560–2572. doi:10.1109/TGRS.2007.900993

Urkowitz, H. (1967). Energy detection of unknown deterministic signals. *Proceedings of the IEEE, 55*(4), 523–531. doi:10.1109/PROC.1967.5573

U.S. Army. (n.d.) *PdM counter explosive hazard: Countermine systems.* Picatinny Arsenal, NJ: Project Manager Close Combat Systems, SFAE-AMO-CCS. Retrieved from http://www.pica.army.mil/pmccs/pmcountermine/CounterMineSys.html#nogo02

Endnotes

[1] PMs must determine what constitutes a “good” performance. For some systems operating in some scenarios, P_d = 0.90 is considered “good,” since only 10 FNs out of 100 true threats is considered an acceptable risk. In other cases, P_d = 0.90 is not acceptable. Appropriately setting a system’s capability requirements calls for a frank assessment of the likelihood and consequences of FNs versus FPs and is beyond the scope of this article.

[2] Statistical tests can determine whether the system’s value is significantly different from the perfect value or the capability requirement (Fleiss, Levin, & Paik, 2013).

[3] Ibid.

[4] Ibid.

[5] Ibid.

[6] Conversely, when Prev is *high*, threat detection systems often exhibit poor values for *N*PV, even while exhibiting excellent values for P_d, P_{fa}, and PPV. Such cases are not discussed in this article, since fewer scenarios in the DoD and DHS involve a *high* prevalence of threat among clutter.

[7] PMs must decide whether the 10 FNs from the first system are acceptable with respect to the tiered system's capability requirements, since the first system's FNs will not have the opportunity to pass through the second system and be found. Setting capability requirements is beyond the scope of this article.

[8] PMs must determine what constitutes a "good" performance when setting the capability requirements for the tiered system.

[9] Statistical tests can show which differences are statistically significant (Fleiss et al., 2013), while subject matter expertise can determine which differences are operationally significant.

[10] Once again, PMs must determine what constitutes a "good" performance when setting the capability requirements for the tiered system.

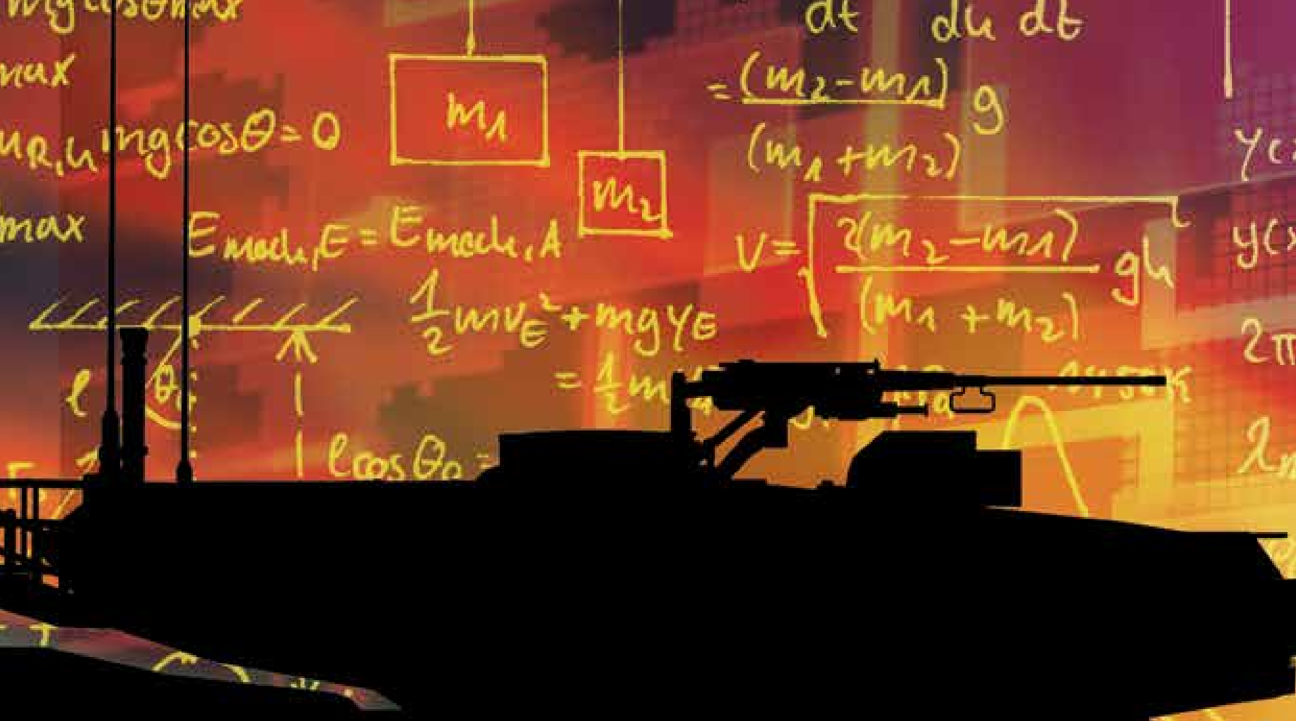
[11] Once again, statistical tests can show which differences are statistically significant (Fleiss et al., 2013), while subject matter expertise can determine which differences are operationally significant.

[12] All four of these metrics are correlated, since all four metrics depend on the system's threshold for alarm. For example, tuning a system to lower its alarm threshold will increase its P_d at the cost of also increasing its P_{fa}. Thus, P_d cannot be considered in the absence of P_{fa} and vice versa. To examine this correlation, P_d and P_{fa} are often plotted against each other while the system's alarm threshold is systematically varied, creating a Receiver-Operating Characteristic curve (Urkowitz, 1967). Similarly, lowering the system's alarm threshold will also affect its PPV. To explore the correlation between P_d and PPV, these metrics can also be plotted against each other while the system's alarm threshold is systematically varied in order to form a Precision-Recall curve (Powers, 2011). (Note that PPV and P_d are often referred to as Precision and Recall, respectively, in the information retrieval community [Powers, 2011]. Also, P_d and P_{fa} are often referred to as Sensitivity and One Minus Specificity, respectively, in the medical community [Altman & Bland, 1994a].) Furthermore, although P_d and P_{fa} do not depend upon Prev, PPV and NPV do. Therefore, PMs must take Prev into account when setting and testing system requirements based on PPV and NPV. Such considerations can be done in a cost-effective way by designing the test to have an artificial prevalence of 0.5 and then calculating PPV and NPV from the P_d and P_{fa} values calculated during the test and the more realistic Prev value estimated for operational settings (Altman & Bland, 1994b).

Biography

Dr. Shelley M. Cazares is a research staff member at the Institute for Defense Analyses (IDA). Her research involves machine learning and physiology to reduce collateral damage in the military theater. Before IDA, she was a principal research scientist at Boston Scientific Corporation, where she designed algorithms to diagnose and treat cardiac dysfunction with implantable medical devices. She earned her BS from MIT in EECS and PhD from Oxford in Engineering Science.

(E-mail address: scazares@ida.org)



INCREASING ARMY SUPPLY CHAIN PERFORMANCE Using an Integrated End-to-End METRICS SYSTEM

Fan T. Tseng, Laird Burns, James T. Simpson, and David Berkowitz

Army Materiel Command and the University of Alabama in Huntsville partnered to develop an integrated end-to-end performance metrics system. The integration includes data pulls from multiple data systems into a metrics calculation and aggregation system that generates strategic performance metrics such as Customer Wait Time (CWT), with capabilities spanning from bottom-up supply chain performance aggregation capabilities to in-depth traceability to source (tactical) level data and documents. To support the best national defense, we must ensure that our warfighters receive the supply support they need in a timely and efficient manner. Supporting this

effort requires a near-real-time system that measures and reports on supply chain strategic performance characteristics such as CWT. Data integrity is an integral part of the process, as is reaching common agreement on appropriate data sources, algorithms to calculate metrics, and the design of a visual dashboard that supports leadership decisions and performance evaluation, with drill-down capability for lower level decision making.

DOI: *http://dx.doi.org/10.22594/dau.15-742.24.01*
Keywords: *performance measurement, metrics, end-to-end supply chain, integration*

Image designed by Diane Fleischer

The current environment for the DoD is one of changing demands: recent high demand for parts needed during engagement in theater has been replaced with inventory rationalization due to drawdowns, new budgetary pressures to reduce costs and expenditures, and a call to emulate commercial practices of efficient supply chain management. In order to achieve greater management visibility across their supply chains, Army Materiel Command (AMC), headquartered in Huntsville, Alabama, partnered with University of Alabama–Huntsville (UAH) in a multiyear project to develop a supply chain performance metrics framework that provides a more comprehensive and integrated end-to-end supply chain metrics system. This research project, named the Enterprise Supply Chain Analysis & Logistics Engine (eSCALE) project, was managed by AMC for the purpose of expanding its scope to all AMC commands.

A principal objective of the project was to create a metrics framework and a prototype analytical calculation and aggregation engine to drive a near-real-time tool (Dashboard) for integrated, supply chain performance visualization, with supportive data reporting. The development of an operational metrics system embedded in a software program was specifically not part of the task. However, the development of a performance measurement framework that demonstrates "thought leadership" and advances in supply chain performance metrics necessitated that we create a prototype system to demonstrate the integration and comprehensive features of such a metrics framework. This project was to look beyond and add to the metrics framework already adopted by DoD.

DoD has previously adopted the supply chain operations reference (SCOR) model (Supply Chain Council, 2016), but AMC wanted a fresh and independent assessment of a metrics framework to aid in redefining existing metrics, provide increased visibility, and enable proactive action in both supply chain and acquisition processes. Our team was free to use guidelines from SCOR and other metrics frameworks, but we were not limited to those or necessarily required to incorporate those views in this research project. The research team was focused on developing a new approach to integrated

supply chain metrics. In particular, we were specifically informed that our task was "thought leadership" on the subject, and not to write executable code for direct implementation.

As the three-phase project progressed, it focused on three dimensions for metrics: *on-time performance, efficiency, and quality*. On-time performance measures whether soldiers in the field received items ordered within a standard lead time for the item. Efficiency focuses on the total cost to the Army for providing the item. Quality focuses on initial quality and fielded life—or reliability—of the item, as well as related quality costs, according to standards for initial quality and field reliability.

The first phase of the project involved identifying and defining metrics that measure supply chain performance of the entire Class IX Army supply chain. Class IX items are repair parts and components for all maintainable equipment, ranging from small parts to complex engines and transmissions for tanks and aircraft. The second phase of the project integrated the metrics into a web-based visualization tool—the "Dashboard." In the third phase, the project sought to develop a framework of a broad-based metrics system to

- measure and evaluate supply chain effectiveness;
- communicate metrics performance standards;
- direct attention to performance areas requiring management intervention;
- identify the root causes of existing or potential performance throughout the supply chain; and to
- more effectively manage total ownership cost.

Project success criteria were not spelled out in a formal requirements document. The sponsoring lieutenant general was seeking not a specific prescribed outcome but a fresh look at devising a metrics framework of highly integrated and near-real-time metrics. With such a framework, senior leadership could effectively assess areas of strong supply chain performance and areas that need attention to improve results. A dashboard

and drill-down capability were seen as desirable characteristics of such a system, but how to derive the functionality and what it would look like were left to the collaborative work between the research team and Army personnel supporting the project.

Measurement Challenges

Measurement systems should be evaluated as a closed-loop control scheme, where multi-organizational-level and multi-echelon metrics constitute an integrated system with performance specifications and feedback loops for ongoing and dynamic performance measurement (Bititci, Carrie, & McDevitt, 1997; Melnyk, Stewart, & Swink, 2004). The metrics need to be aligned with internal organizational processes and external customer and supplier performance using common definitions, data sources, and calculations, while balancing the tension between conflicting organizational priorities (Melnyk et al., 2005). Many organizations that seek to improve

their performance measurement system have initially evolved to function-specific and isolated, or stovepipe, metrics, but now want to implement a more integrated metrics system.

For instance, Figure 1a (upper left corner) displays the existing metrics paradigm when the eSCALE project was initiated in 2008. The metrics, for the most part, existed in a stovepipe environment, and their definition, source data, and interpretation varied within different AMC subordinate commands. Metrics were collected and reported, and did not reflect the performance of the entire supply chain. Moreover, the use of averages often hid issues that affect performance (Oliver, Delbridge, & Lowe, 1996)

This presented an opportunity to develop a standardized method of pulling and analyzing data across the command from multiple data systems to provide a true enterprise-wide view of the AMC's supply chain, even when the existing data systems were unable to communicate across information systems.

The new approach presented in the eSCALE project, as illustrated in Figure 1b, 1c, and 1d, involves moving away from averaging performance across the subordinate commands to measuring and monitoring exceptions in each area of the supply chain. This approach, combined with a near-real-time data engine, provides the ability to identify and address issues at the tactical level that, for example, would show overstock or understock levels by location, or production lead times beyond acceptable targets. Additionally, it provides the capability to drill down to the individual item level to identify and implement operational performance improvements. This approach identified issues with particular National Item Identification Numbers (NIIN) that could be readily targeted for performance improvement, such as the percentage of NIINs with administrative lead time (ALT) or production lead time (PLT) that were beyond acceptable standards, or stocking locations with inventory levels above requirements.

One of the key contributions beyond the SCOR approach was the selection of Customer Wait Time (CWT) for the warfighter (Retail CWT) as a focal metric for on-time performance. This is a departure from more traditional supply chain performance metrics, such as parts availability or inventory turns, and is similar to a supply-chain-spanning version of order-to-delivery (ODT) time, where we also apply CWT at each upstream echelon across four or more echelons in the supply chain (Keebler, Manrodt, Durtsche, & Ledyard, 1999). Retail CWT serves as an objective surrogate for readiness by assessing on-time performance of supply delivery to warfighters regardless

of their location in the world. CWT for upstream echelons includes real-time indicators, before delivery of the end item to the warfighter, on whether prior processes are performing on time.

To support the strategic-level metric of CWT for each echelon of the supply chain, a suite of metrics was developed for pipeline, supply, and demand management with tools for detailed analyses (Figure 1b). This approach provided leading indicators of potential issues that would, if not addressed, extend Retail CWT for the warfighter beyond targeted performance. These metrics include measures of key processes *earlier* in the upstream supply chain that would give supply chain managers early warning signals that they need to take some action *before* they fail to deliver on time to the warfighter. For instance, supplier delivery performance measures are developed to serve as leading indicators of Supplier CWT to the wholesale supply chain that eventually affects Retail CWT for the warfighter. In other words, performance measurement of upstream echelons such as suppliers and wholesale distribution helps predict expected Retail CWT for a future time. When the leading indicator falls out of the range of tolerance, it will be flagged to signal the need for some action.

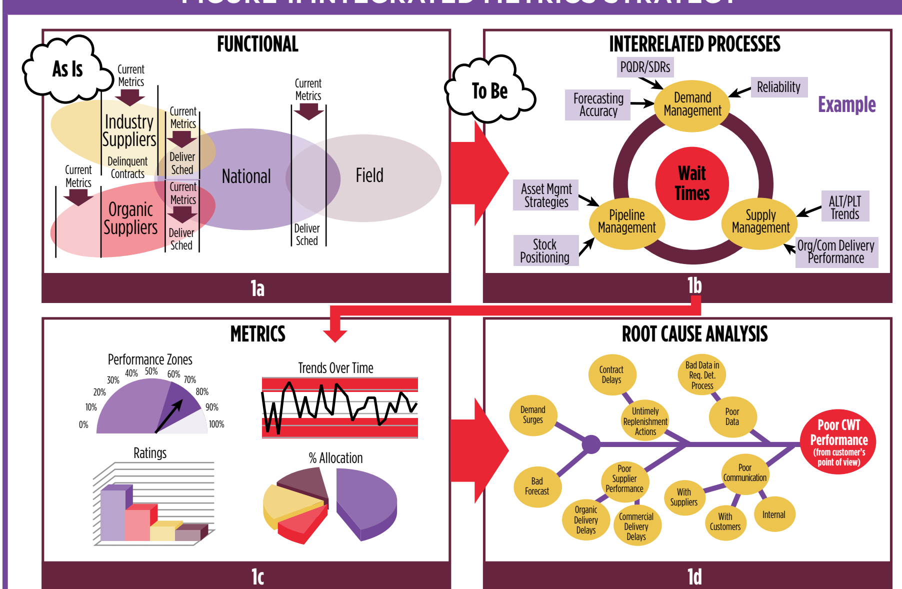
This approach improves the current process, where AMC compiles metrics that identify the number of days the soldier waits for the item, without the ability to drill down for issue investigation. The new approach also allows monitoring of drivers of CWT at each upstream echelon that could extend Retail CWT, such as demand variation, increased lead time, bad data, poor stock positioning, delivery delays, or other root causes of increased CWT (Figure 1d). Moreover, existing metrics do not allow for both a comprehensive view within a command and an aggregated view across commands along different dimensions, such as by weapon system or Army Class of Supply for parts.

Another modification of the SCOR framework dealt with the reliability metric. Reliability in the SCOR model is defined as "the predictability of the outcome of a process" (Supply Chain Council, 2016, p. 2.1.1), but does not apply to actual quality or reliability of the product or system delivered, as SCOR loosely refers to a product in "perfect condition" (Supply Chain Council, 2016, p. 2.1.1). In SCOR, reliability includes Perfect Order Fulfillment, comprising five standards that include perfect quality—which primarily focuses on perfect order documentation and the product's arrival in "good" condition. Yet in complex products and systems such as those deployed by the Army, reliable transactional processes and accurate paperwork are inadequate to ensure product reliability. In the eSCALE project,

reliability was defined by the standard adopted by DoD, which specifies that reliability is the failure-free performance of a system or product for a specified time under specified conditions (DoD, 2009; Information Technology Association of America, 2008). Reliability in this project includes initial quality and fielded reliability, using the definition above. A detailed review of the quality and reliability aspects of the eSCALE is beyond the scope of this article and is detailed elsewhere (Burns & Nicholls, 2013).

The objective of the eSCALE project was to provide visibility across the supply chain with performance measures that afford an integrated, consistent, and near-real-time view of strategic, tactical, and operational measures. With the right measures and analytical capabilities, individual supply chain managers will be empowered with the information necessary to identify and resolve issues, and make informed resource allocation decisions. Improved visibility, for example, supports improvements in inventory turns, inventory investments, and customer fill rates to soldiers. Aligning appropriate information with the existing organizational structure is expected to increase visibility across organizational boundaries. This will help improve cooperative and collaborative efforts to enhance the supply chain since increased transparency of information becomes an instrument of change rather than a source of conflict within the supply chain.

FIGURE 1. INTEGRATED METRICS STRATEGY



Integration with Acquisition

The acquisition process is integral to the new integrated supply chain performance measurement system. The new system establishes tolerances and performance norms for the suppliers, reports actual performance, and highlights exceptions to acceptable performance. As the system monitors the supplier's performance on established metrics, AMC will have the capability to create a scorecard for each organic and industrial supplier. The scorecards can aggregate performance by Life Cycle Management Command, weapon system, supplier, or from other perspectives for senior leadership review, and they provide drill-down capability to the individual item and contract level, as well as the aggregated contractor report card.

It has been known for many years that integrating sourcing performance, such as via acquisition price analysis, can have a considerable effect on overall supply chain performance (Beamon, 1998).

On-time delivery performance may suffer due to forecasting error or unexpected demand changes, which are highlighted in the metrics system. But other issues with poor CWT performance may arise. The drill-down capability of the system allows the item manager to identify issues that lead to CWT violations. Where late deliveries are not caused by forecasting error or unexpected demand changes, the metric system highlights in near-real-time other causal issues so they can either be corrected by the supplier or contract modifications can be implemented before the poor CWT performance reaches the warfighter.

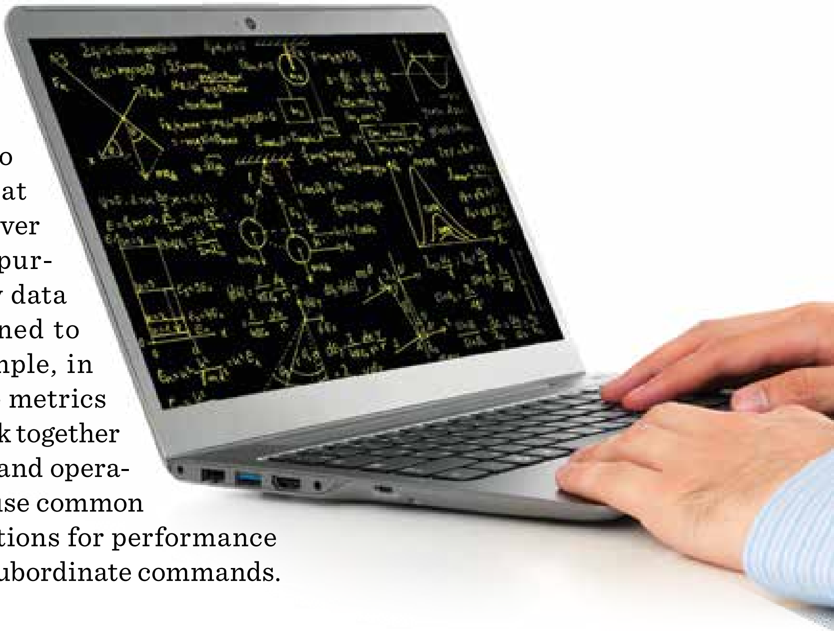
It has been known for many years that integrating sourcing performance, such as via acquisition price analysis, can have a considerable effect on overall supply chain performance (Beamon, 1998). Excessive CWT can reduce readiness and increase costs by causing a need for higher safety stock inventory. Quality and fielded reliability issues also increase cost by requiring more items to be provided to achieve an operational tempo, while also increasing maintenance and transportation costs. Each of these cost effects can be included in performance and pricing analysis at acquisition, and for some suppliers both may be present at the same time. For acquisition evaluation, the eSCALE Dashboard complements, but does not incorporate,

lagging performance data from the Contractor Performance Assessment Reporting System (2016), Past Performance Information Retrieval System (2016), or Governmental-Industry Data Exchange Program (2016).

Suppliers whose performance consistently exceeds CWT specifications can be evaluated based on actual delivered costs that are higher than quoted costs (Burns & Nicholls, 2013), providing an accurate comparison of supplier performance. For specific supplier performance, integrating these costs across AMC commands rather than by separate contracts with the supplier can show aggregated performance over time. However, for full implementation, data challenges must be addressed in the process.

Data Challenges

There are many Army data systems that handle the complex transactions and processes necessary to support our warfighters. Yet the many Army supply chain data systems are not linked into a single integrated system with highly functional connectivity to the other systems, making comparisons of performance and identification of improvements difficult (Siegl, 2008). The implementation of the Army Logistics Modernization Program (LMP), an enterprise resource planning system, for example, has created a more integrated planning system but has not facilitated an integrated, near-real-time, performance measurement system for the Army-wide supply chain. One of the reasons is that multiple data systems are required for an integrated near-real-time system to be effective, and similarly to commercial systems that have been implemented over time to serve different purposes, many of the Army data systems were not designed to work together. For example, in an integrated system the metrics should be designed to work together at the strategic, tactical, and operational levels, and should use common terminology and calculations for performance measurement across all subordinate commands.



In the development of the eSCALE performance metric system, the research team integrated data inputs from more than 10 different, nonintegrated data sources into a single performance metric architecture. The data sources, records, and fields used for each of the calculations were developed by the research team and approved by the Army sponsor with concurrence from Army subject matter experts (SMEs). This avoided incomplete data sets from one source, while providing a single Army-authorized data source for each of the calculations.

Delivered Results

By the end of the project, the architecture and framework for a full, on-time, performance metrics suite had been developed and delivered, along with a corresponding integrated framework for an efficiency metrics suite. The architecture for the quality and reliability framework had also been developed, but had not been prototyped and integrated with the on-time performance and efficiency metrics due to funding reduction as a result of sequestration. More discussion on these results and the subsequent development of the quality and reliability metrics appears later in this article.

Literature Review

The scale and complexity of Army supply chains necessitate a large and complex organization to oversee them. Over decades, organizational units with responsibility for various activities have been created, merged, and in some cases dissolved, leading to a supply chain governed by historical evolution and layers of regulation, policy, and directives. Each subordinate command specializes in supply chain management activities specific to particular weapon systems, so not surprisingly, many of the individual organizational units focus very specifically on their own domain of responsibility.

As a result of the need for specialization, over many years this separation of command responsibilities has resulted in a set of practices that can be viewed as "silos of excellence." One might wonder: would a consolidated distribution agency such as Defense Logistics Agency (DLA) solve the silo challenge? Not likely, as the solution is not that simple. While DLA has responsibility for sourcing and distributing many national stock numbers that flow through AMC, it also interacts with many suppliers, technical engineering activities, and customer groups. Also, many of the issues arise from complex and long-lead-time systems such as engines and transmissions, which require significant technical support over time. From this, our conclusion would be that the metrics framework may be of considerable use to agencies such as DLA. The multi-echelon integration across diverse data platforms is one of the strengths of the metrics framework and visual dashboard. Further benefits of this will be explained in detail in the Methodology section.

Due to the silos of functional focus that have evolved over many years, it is often difficult for management to obtain a consistent view or a comprehensive assessment of overall supply chain performance due to variation in data sources, definitions, and/or interpretation. There are many performance measurement models in use, some widely used and others that are organization-specific (Bititci, Carrie, & Turner, 2008). One widely adopted model is the SCOR model. The SCOR model helps organizations to focus performance measurement on five key processes across supply chains: plan, source, make, deliver, and return (Siegl, 2008). Another popular model is the balanced scorecard developed by Kaplan and Norton (Kaplan & Norton, 1992), which links performance measures from the financial, customer, internal business process, and innovation and learning perspectives.

Integrating and updating key business processes, avoiding inefficiencies and resource conflict across the supply chain, can be even more challenging when diverse organizations develop their own unique objectives and corresponding metrics (Henderson, 1994). This required a strategic rethinking of the entire approach to AMC enterprise-wide performance metrics as an integrated system. One way to view this approach was explained by Francis (2001), who argued for increasing operational agility to improve performance, using the example of lessons learned from military agility in the field. While extended lead times and other physical supply chain constraints make it challenging to achieve agility with all Class IX parts, lessons can be applied for those items that specifically require higher degrees of agility.

The development of an integrated suite of supply chain metrics should include strategic measures of on-time performance, quality performance, and cost performance. A metrics suite should also include drill-down subordinate metrics to support tactical- and operational-level decisions (Gunasekaran, Patel, & Ronald, 2004). This integrated cross-organizational approach supports the Joint Supply Chain Architecture (JSCA) initiative (Fletcher, 2011), which has adopted the CWT metric to measure speed of delivery to warfighters (DoD, 2012, Chapter 15.4).

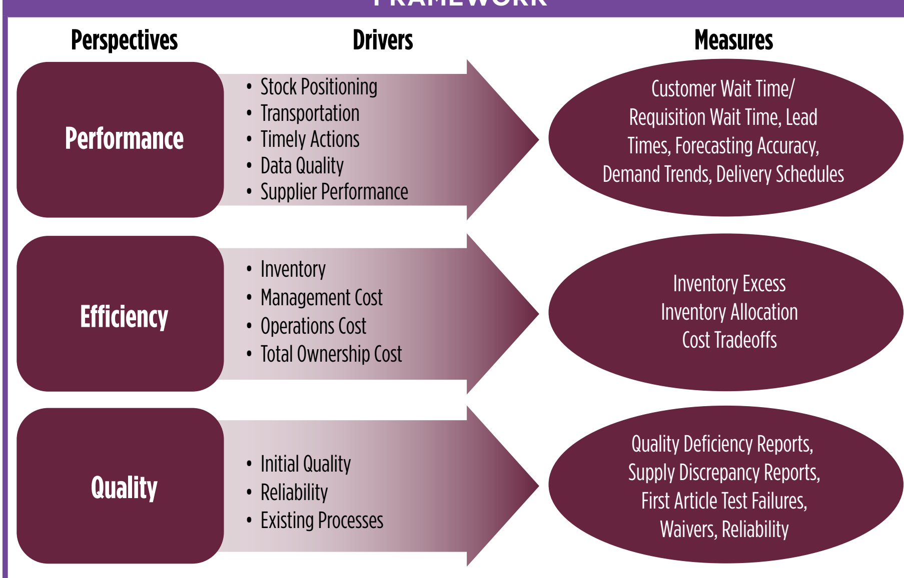
Methodology

The methodology for proposing a solution was selected from a variety of options. The first step for the Army was to choose a partner with significant supply chain expertise and knowledge of Army's supply chains, but also one that has the ability to bring diverse, unbiased perspectives and solutions to existing requirements. A UAH team of academic faculty, research scientists, and graduate students from fields such as operations research, operations management, supply chain management, marketing, economics, information systems, systems engineering, reliability, and modeling and simulation was selected to support the project. To shorten the learning curve, AMC scheduled joint-learning seminars with the UAH research team and AMC supply chain managers, supply chain specialists, and representatives from subordinate commands and organizations, such as Integrated Material Management Center, Security Assistance Management Directorate, Logistics Support Activity, Rock Island Arsenal, Corpus Christi Army Depot, and AMC G6 and G4 to discuss the overall approach and unique aspects of the Army's supply chain. SCOR principles were quickly determined to be a key element in view of DoD's adoption of the framework. It was specified that the solution would need to utilize existing data sources

rather than create requirements for new data sources. This task was complicated by the necessity to accommodate AMC's recent transition to a new SAP-based information system known as LMP (Logistics Modernization Program), which at the beginning of this multiyear project was a new deployment at Aviation and Missile Command (AMCOM). To aid in this task, potential solutions were discussed and analyzed during frequent brainstorming sessions between UAH and government representatives.

The UAH team understood the significant differences in the Army supply chains, which are geared towards rapid deployment, sustainment, and drawdown for wartime activities. This is different from commercial organizations, which tend to focus on market-based metrics such as profits, return on investment (ROI), and stock prices, and often experience considerably smaller variations over time in the scale of operations. To bridge this difference, UAH and AMC officials conducted phone interviews and site visits with several multinational commercial organizations known to apply best practice supply chain strategies. The initial seminars, meetings, and site visits resulted in an agreed-upon solution framework, which is represented in Figure 2. This article focuses primarily on CWT as an example and primary metric for on-time performance, but all of the metrics in Figure 2—and more—were included in the eSCALE performance metrics system. Space limitations prohibit a full review of all of the elements of the metrics project, which when documented filled a book.

FIGURE 2. PERSPECTIVES OF THE eSCALE MEASUREMENT FRAMEWORK



Solution Perspective and Framework

A major objective in developing the end-to-end metrics suite was to adopt a comprehensive and multidimensional approach to measure and evaluate overall supply chain performance along the dimensions of *on-time performance, efficiency, and quality*. *On-time performance* of the supply chain is driven by the chain's ability to deliver to the end customer on a time standard, specifically focusing on *speed*. Based on this notion, the approach in developing on-time performance metrics places CWT as a key metric, with the rest of the suite consisting of metrics that are likely to affect the key metric of CWT. Accordingly, the on-time performance metrics suite provides a framework for insightful analysis of how on-time performance is impacted by key drivers such as stock positioning, transportation, timely action by administrators and suppliers, data quality, and supplier delivery performance. In other words, CWT can be affected by factors such as where the stock is positioned in the supply chain, the speed with which the item was shipped to other members of the supply chain, and the timeliness of the procurement decisions made by the item manager. To assess the impact of on-time performance drivers, we measure and report metrics such as requisition wait time (RWT), demand variation, stock position, and delivery schedules, as detailed in the next section.

The Army's increased attention to inventory levels has elevated the importance of the *efficiency* metric. A recent Government Accountability Office study (Solis, 2009) reported an annual average of about $16.3 billion in secondary inventory, of which $3.6 billion (22%) exceeded current requirements. The study also revealed that, while the Army experienced excess secondary inventory, the Army had annual inventory deficits based on requirements of about $3.5 billion per year. The lack of alignment of existing inventory and current requirements was attributed to a lack of cost-efficiency metrics and goals, and inaccurate demand forecasting.

Our research revealed that inventory management, management costs, operations costs, and total ownership costs (General Accounting Office, 2003) are the primary drivers of efficiency in the Army supply chain. Metrics such as inventory excess, inventory allocation, and cost tradeoffs (e.g., repair and procurement) were developed and measured to assess the status of the drivers of supply chain efficiency.

Metrics for measuring *quality* seek to determine whether assets are functioning and performing in the field as designed and manufactured. The key drivers are initial quality, reliability, and performance of existing processes. Initially, the analysis focuses on supply deficiency reports and quality

deficiency reports. Additionally, first article test failures and waivers can serve as metrics to assess the key drivers. It was necessary in eSCALE, and in a follow-on research project, to develop algorithms to measure total costs for fielded reliability of secondary components based on planned versus actual life, which is important to containing life-cycle costs (Burns & Nicholls, 2013).

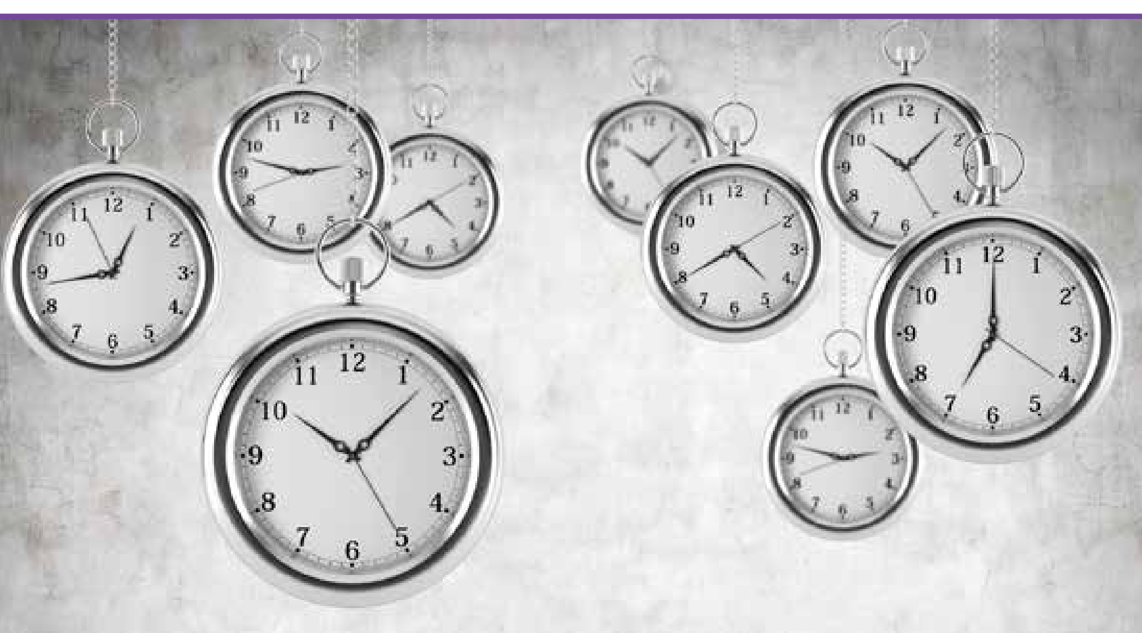
Once the initial key metrics have been formulated, the focus shifts to designing a performance measurement system that will transition away from averaging across the entire inventory to defining acceptable levels of performance and subsequently counting the number of "bad situations" (i.e., flagged issues) for each individual part that does not meet the defined standards for on-time performance, efficiency, or quality. This strategy is common across highly successful commercial organizations.

The advantage of this approach is that each individual transaction, such as a part requisition, is evaluated against the standard prior to aggregating the counts to a metric. The counts can be aggregated from multiple perspectives, such as by supply support activities (SSA), weapon system, region, priority,

and supplier, as well as by key locations in the supply chain, to show where the greatest concentration of supply chain improvement opportunities exists. The next section will detail the specific approaches used to develop the on-time performance component of the project.

The architecture of the metrics framework was designed for multiple stakeholders. For item managers, the framework identifies performance metrics and issues that will help them manage and execute their tasks for NIINs for which they are responsible. This is provided in the drill-down capability to the item level. For senior leadership and for mid-level managers, the roll-up capability aggregates and summarizes performance by responsible branches, divisions and commands, as well as by weapon system, organic or industrial supplier, and other categories of review. For issues that merit additional attention, the drill-down capability can identify areas of strong performance and areas needing additional attention at multiple levels of scrutiny, at the leader or manager's discretion.

During the development and prototyping of the metrics, we had extensive involvement with item managers and mid-level users, ongoing periodic reviews with senior leadership to obtain feedback, and confirmation of definitions and calculations for metrics and for the design of a dashboard metrics system for their use. An example of this process is detailed in the next section.



One challenge that we addressed in the Dashboard design is the issue of trade-offs of on-time performance, efficiency, and quality. A well-designed metrics system identifies and highlights such issues for decision makers. One advantage in the Army is that the hierarchical command structure naturally allows for issues regarding performance trade-offs to be identified to leadership, where well-informed leaders can make decisions on the trade-offs that cannot be encoded into any particular metrics system. We saw our role as one of helping to inform users, managers, and senior leadership of areas of high and of acceptable performance, while also highlighting areas where additional attention can lead to increased performance.

On-Time Performance

Recall that the objective of this research was to develop a supply chain system that maximizes overall performance of the supply chain and the related acquisition processes. This study proposed that on-time performance, efficiency, and quality are key metrics for assessing the overall performance of the Army supply chain. The focus of this section is on-time performance, which is driven by the ability of the supply chain to deliver to the end customer on a time standard. This details the major steps adopted in the development process of the eSCALE system that measure, report, and analyze on-time performance of the supply chain. They include the following:

- Task—Create a new effective way to measure enterprise supply chain performance.
- Concept—Unmask and trace issues, and predict trends.
- Development of Logic—Define tolerances to create standardized metrics.
- Data—Identify data sources, and acquire and verify data.
- Proof of Concept—Test logic against the inventory to get actual counts.
- Visualization—Create a visualization tool (Dashboard) to display results.
- Validation and Documentation—Assure data and process consistency and reliability across commands, and prepare for implementation.

Task. The existing metrics program's "stovepipe" environment, variations in definitions and data availability, and dependence on the use of averages preclude an accurate assessment of the performance of the enterprise-wide supply chain. Hence, the task for this project was to improve the existing system by offering the Army an integrated approach to measuring enterprise supply chain performance.

Concept. Given the prominence of readiness as the definitive measure of supply chain effectiveness, our concept for addressing the on-time performance metric focused on CWT as a viable surrogate for readiness. More specifically, we count actions, rather than report averages, which potentially mask performance issues. The goal is to know what percentage of the NIINs do not meet established performance standards. Reporting that percentage is more informative since the total number can be misleading, given the variation in size of organizations. Of course we must count both total NIINs and NIINs with issues to determine the performance. Figure 3 describes the counting logic process. The data for NIINs is evaluated against Office of the Secretary of Defense (OSD) standards. For each NIIN, if the preset standard is being violated, a 1 (yes) is assigned and an issue count is generated. Alternatively, a 0 (no) is assigned if the performance is within the standard. These results (0 and 1) are rolled up at higher levels to reflect tolerance bands and trends, as will be shown later in the discussion of the Dashboard.

FIGURE 3. COUNTING METHODOLOGY: LOGIC EXAMPLE

LOGIC EXAMPLE

The count standard assesses raw data against an acceptable standard. When the data are outside the standard, a bad 'count' is generated. These are rolled up at higher levels to reflect tolerance bands and trends.

Development of Logic. The next step in the process is the development of the logic underlying the system. This step defines the tolerances to create standardized metrics across the commands. Standards were established for CWT and measures of major factors affecting CWT, including ALT and PLT, data discrepancies, demand quantity variation, delivery performance for new buys, maintenance and operations (M&O) delivery performance for both commercial and organic/depot suppliers, stock positioning, and RWT.

Supply chain managers make decisions about item deliveries based on the contract delivery schedule (e.g., deliver items within 60 days of contract signing). Yet, issues occur when the LMP data do not accurately portray the correct contract signing date. This issue is captured in the metric of ALT, which is the time between the date the item manager requests the item and the date the contract with the delivery schedule is signed. The tolerance metric for ALT states

- *If the deviation of average ALT is greater than 25% from the LMP ALT in the last 12 months, logic denotes a 1, signaling an issue.*

The numbers included in the logic above (25% and 12 months) and thereafter are proposed as examples and should be determined by proper decision-making processes for the respective NIINs.

Demand variation attempts to capture the extent to which the current demand is either the highest or lowest value compared to the previous 12 months. Tolerances might include metrics such as

- *If (the current) demand is greater than monthly demand for each of the recent X months, the logic denotes a 1, signaling an issue (a change in demand).*
- *If (the current) demand is less than monthly demand for each of the recent X months, the logic denotes a 1, signaling an issue.*
- *If demand is trending upward or downward for Y consecutive months, the logic denotes a 1, signaling an issue.*

Late deliveries are captured in the delivery tolerances for new buys, for both commercial and organic M&O deliveries. Examples of tolerance statements include

- *If the percentage of late deliveries for the current month is greater than 10%, logic denotes a 1, signaling an issue.*
- *If the percentage of cumulative late deliveries is greater than 10%, logic denotes a 1, signaling an issue.*

Using the flags by NIIN (where a signal = 1), item managers can quickly identify exceptions requiring attention. Inventory issues are related to how item managers administer their contract and procurement plans. The item manager's contract and procurement plan reflects how the item manager plans to fulfill the requirement for a particular item or component beyond the scheduled due-ins. The due-ins plus the items programmed in the contract and procurement plan should equal the forecasted requirements. The eSCALE system identifies data gaps between the demand forecast and the total supply expected from the due-in and the contract and procurement plan. The tolerance statement is

- *If the difference between the total of all due-ins, contract and procurement plans, and the forecasted requirements is more than Z units in month 1 of the planning horizon, the logic denotes a 1, signaling an issue.*

Since serviceable items should always be located as close as possible to the customer and repairable items should be located at the appropriate repair facility, a major goal for the new system is to identify where inventory might be better positioned in the pipeline. For instance, a serviceable depot-level repairable item located in an organic repair facility rather than the SSA where it could be issued to the customer would be considered an issue. Moreover, an unserviceable item that can be repaired only by the contractor should be located at the contractor repair facility rather than the organic repair facility, the SSA, or an area operated depot (AOD). Figure 4 demonstrates how the eSCALE system identifies stock locations. A flag of 1 represents inventory that can be better positioned.

FIGURE 4. PERFORMANCE PERSPECTIVE: LOGIC STATEMENTS

Stock Positioning: Part Type
Identifying where inventory position can be improved in the pipeline

	Contractor Repair Facility	Organic Repair Facility	SSA	AOD/CCP
DLR Serviceable	1	1	0	0
DLR Unserviceable	0	0	1	1
FLR Unserviceable	1	0	1	1
Contractor Repair Only Unserviceable	0	1	1	1
Contractor Repair Only Serviceable	1	1	0	0

Note. DLR = Depot Level Reparable; FLR = Field Level Reparable.

A more unusual stock-positioning challenge relates to stock positioning at the SSA. One challenge occurs when there is unforeseen demand for an item not on the Authorized Stock List (ASL). The second challenge occurs when items continue to exist on the ASL (Requisition Objective > 0) when there is no current demand for the item. The tolerance statements for these two conditions follow:

- *If the Requirement Objective (RO) = 0 and demand is greater than 3 in the past year, logic denotes a 1, signaling an issue.*

- *If the Requirement Objective (RO) is greater than 0, and demand = 0 in the past year, logic denotes a 1, signaling an issue.*

The amount of time that the SSA waits to receive the requisitioned item from the wholesale organization, such as AMCOM, is RWT. RWT is also an important factor in determining the amount of time the final customer must wait for an item or component, CWT. Hence the following tolerance was established:

- *If the requisition wait time is longer than TRANSCOM [Transportation Command] regulation for CONUS/OCONUS [Continental U.S./Outside Continental U.S.], the logic denotes a 1, signaling an issue.*

CWT is the seminal variable in the eSCALE system, since it is our surrogate for readiness. The variables used in the model discussed above (e.g., demand variation, stock positioning) were selected because of their potential negative impact on CWT. The customer is the warfighter, who secures the product/item/component from the SSA. The ordering process begins when the soldier completes the appropriate order paperwork. The resulting requisition records the date and time the requisition was filed at the SSA. The process is completed when the soldier receives the item from the SSA. The time between the date and time of the requisition submission and the date and time the soldier received the item is the CWT. While most stock items are available for immediate distribution to the customer, the eSCALE system identifies instances when the CWT exceeds the established tolerances.

- *If the Requisition Objective (RO) is greater than 0 and more than 15% of the requisitions for an item took longer than 3 days, logic denotes a 1, signaling an issue.*

- *If the Requisition Objective (RO) = 0 and more than 15% of the requisitions for an item took longer than TRANSCOM standards, logic denotes a 1, signaling an issue.*

The metrics discussed above were developed out of many meetings and discussions with different Army agencies across the supply chain. For metrics to be useful, they must be supported by data that are accessible, accurate, consistent, and current.

Data and Proof of Concept. The development of metrics and the logic for metrics calculations and flags for performance improvements cannot be separated from the process and effort of identifying data sources, and acquiring and verifying data. The research team collaborated closely with Army SMEs in this effort, with approvals by senior leadership.

Since the Army supply chain consists of organizations that are specialized in their functional areas, it soon became evident that identifying data sources for the metrics was the first essential step. Working groups were formed, with representatives who are mostly SMEs from organizations managing or supporting the supply chain, to reach agreement on metrics and standards for the supply chain Dashboard and to identify data that will be used to calculate metrics and data sources. Once the data sources for metrics were identified, more time-consuming tasks followed our acquisition of sample data for analysis and testing of proposed metrics to prove the concept.

Proof of concept involved testing the logic against the available data and evaluating the logic with Army SMEs. The tasks were not in sequential order—it took a lot of back and forth among the UAH team and Army commands. Iterative steps ensued during this collaborative process, from defining metrics to identifying data sources, acquiring data, verifying

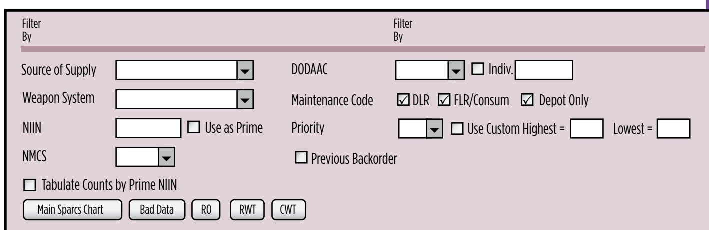
data, testing metrics with acquired sample data, and testing metrics with full-scale data. It was essential to do this in order to obtain a consistent, enterprise-wide view and a comprehensive performance assessment of the supply chain.

Visualization. A new visualization tool, the Dashboard, was developed to facilitate a departure from reporting averages across the entire inventory and focus on counting of "flagged" areas in the individual supply chain elements with the ability to drill down to the individual item level. Figure 5 provides an example of the Dashboard, which includes the standards and logic for the CWT metric. To support the drill-down and root-cause analysis, the Dashboard presents the total number of issues by category for each month and year, and the specific number of issues by potential causes, such as demand variation and stock positioning. Possible sources of issues related to the various causes of CWT are also listed. This top-level chart summarizes all issues across the various categories. This particular chart—showing only notional data for confidentiality reasons—reports at the bottom left a total of 25,070 NIINs for the requested supplier, dates, and maintenance codes. Of the 25,070 NIINs, 5,007 have at least one issue, such as late delivery or a documentation issue. Hence, 20,063 have no (0) issues. A further reading reveals that 3,322 NIINs have one issue and 4 NIINs have five issues. You

will note, that in this particular analysis, most issues occur because of data issues ("Bad Data"), improvable stock positioning ("Stock Positioning"), and excessive time to process requisitions (RWT).

FIGURE 5. PERFORMANCE PERSPECTIVE: DASHBOARD DEVELOPMENT

Available Filter Options



The chart sums up all problems across various categories. For instance, **out of 25,070** NIINs, **4** are flagged as having problems in five categories.

Top Level

Sum of Problems	#NIIN	Bad Data	Demand Var.	New Buys	Commerce	M&O Organic	Inven. Probs.	Stock Pos.	RWT	CWT
9	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
5	4	4	3	0	0	2	0	4	4	4
4	102	91	14	0	0	10	0	101	101	91
3	519	309	26	0	0	27	0	486	475	236
2	1,059	386	27	9	9	17	0	851	673	214
1	3,322	2,201	15	0	0	28	0	643	268	169
0	20,063									
Probs	5,007	2,990	85	0	0	84	0	2,086	1,471	714
All	25,070	18,059	9,699	0	0	2,896	0	9,170	5,972	6,022

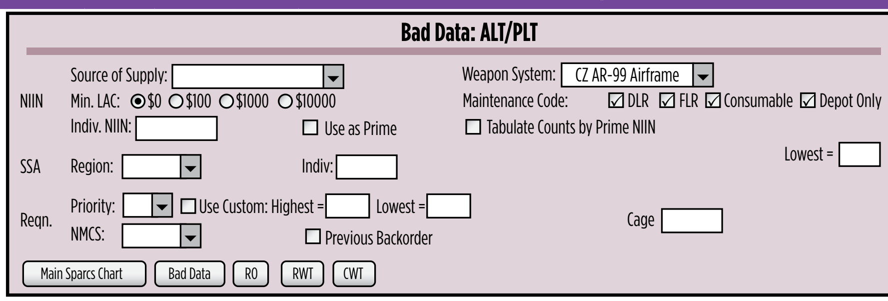
Individual columns illustrate specific metric areas. For M&O Organic, **84** NIINs had greater than 10% delivery delays out of 3,351.

Note. DODAAC = Department of Defense Activity Address Code, Sparcs = supply problem analysis, reporting and categorization system.

Figure 6 represents the Dashboard's drill-down capability, which provides a more detailed presentation of *bad data* issues related to ALT and PLT for the CZ AX-99 Airframe (actual weapon system code disguised). In this notional

example, there are 91 NIINs reported. Seventy-one have at least one bad data issue related to ALT or PLT, so only 20 NIINs have zero (0) ALT or PLT issues. Eighteen NIINs report one issue and 53 report two issues. Of the 18 NIINs with only one issue, 13 are data discrepancies related to ALT and 11 are data discrepancies related to PLT.

FIGURE 6. PERFORMANCE PERSPECTIVE: BAD DATA DRILL DOWN—ALT/PLT



Bad Data

Sum of Problems	#NIIN	ALT	PLT	
2	53	53	53	⇐ NIINs with Bad ALT and Bad PLT Values
1	18	13	11	⇐ NIINs with either Bad ALT or Bad PLT Values
0	20			
Probs	71	66	64	Click for drill down on root cause
All	91	91	91	

Note. LAC = latest acquisition cost, NMCS = not mission capable supply.

To examine additional drill-down capabilities to understand issues, using an example, the item manager simply clicks on the number "13" under ALT in Figure 6 to review details on the 13 NIINs with ALT issues. Figure 7 presents analysis of the 13 NIINs where the actual ALT dates differ from the system's data dates by more than 25%, a preset level of tolerance. This report reveals each of the 13 NIINs with both a distribution and numerical percentage deviation between the actual ALT and the ALT reported in the LMP data files. For instance, the percentage deviations between the actual ALT and the ALT from the data file are shown in the highlighted (and disguised) NIIN on the right (83609549146) to be 41%. Since 41% is greater than the acceptable tolerance of 25%, it is flagged on this screen and the item manager can investigate the issue for that particular NIIN by clicking on the

NIIN for additional drill down. The results for Figure 8 show that the NIIN is for a spherical gear for weapon system CZ. The actual ALT is 28.8 days and the LMP system ALT (days) is 20.4 days (data disguised). The manager should investigate the causes of this discrepancy.

FIGURE 7. PERFORMANCE PERSPECTIVE: BAD DATA DRILL DOWN—ALT/PLT: ALT

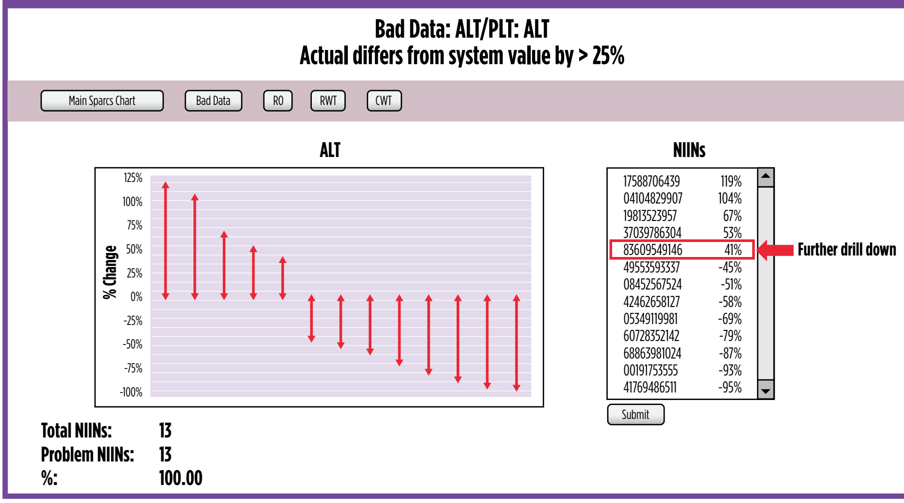


FIGURE 8. PERFORMANCE PERSPECTIVE: BAD DATA DRILL DOWN—ALT/PLT: ALT: SELECTED NIIN

Bad Data: ALT/PLT: ALT: Selected NIIN

Main Sparcs Chart | Bad Data | RO | RWT | CWT

NIIN	**83609549146**
Prime NIIN	**83609549146**
Nomen	**Gear, spherical**
Weapon System	**CZ AR-99 Airframe**
SOS	**AMCOM Aviation Command (B17)**
Actual ALT (days)	**28.8**
System ALT (days)	**20.4**
% Difference	**41.2%**

Note. Nomen = nomenclature, SOS = source of supply.

Validation and Documentation. Before the Dashboard can be implemented, steps must be taken to assure data and process consistency and reliability across commands. Data validation is a major milestone before tool

deployment. Validation means signing off on the data—each data element—being analyzed in the Dashboard, including calculations, data sources, and data analysis. Validation can also include the Dashboard layout and how the data are presented. A data validation plan should be made to determine the validation process, approach, and data teams.

Documentation provides a path to communicate our methodology and to meet Army's information assurance certification requirements. It also helps the data validation process. We documented the project with specifications, data structures, logic algorithms, and user applications and instructions to help educate item managers and other users on the capabilities and interfaces of the performance metrics system framework.

Benefits and Application

The delivered architecture for how to implement the improved metrics framework has allowed the integration of existing metrics, such as those adopted through SCOR, with improvements such as a focus on CWT that is integrated with efficiency. In addition, through continued development of the quality and reliability metrics initiated in this project and further developed in a follow-on project (Burns & Nicholls, 2013), algorithms for assessing the cost of initial quality and fielded reliability have been developed as a complementary metrics architecture for the more fully developed on-time performance and efficiency metrics in the eSCALE project.

The proposed metrics system will enable managers to shift their focus to measuring and evaluating performance, communicating performance expectations, directing attention to areas requiring management intervention, and helping to identify root causes of performance challenges for both AMC and its organic and industrial suppliers. More importantly, it will help expose critical issues and provide the basis for launching new initiatives for improving the supply chain. For instance, following the identification of challenges, opportunities, and goals, a "tiger team" of SMEs could be formed within various segments of the supply chain.

One such example was a recent eSCALE project that conducted simulation studies providing insights on how to improve supply chain efficiency while decreasing the time it takes to deliver long-lead-time items. It was observed that lead times for aviation and missile parts have been increasing, contributing to increased inventory pipeline costs. To reduce lead times and overall inventory pipeline costs, new contracting strategies were needed

for deploying strategic Work-in-Progress (WIP) inventory at different locations in the upstream supply chain. Strategically placed WIP has the potential to increase the supply chain's capability to meet demand surges and meet weapon system readiness requirements while helping to reduce overall inventory. Another example of the increased visibility was the identification of issues to improve supply chain performance in theater and other areas. Soon after the issues were identified, item managers were able to enact strategies that resulted in real supply performance improvements to our warfighters.

> ***The proposed metrics system will enable managers to shift their focus to measuring and evaluating performance, communicating performance expectations, directing attention to areas requiring management intervention, and helping to identify root causes of performance challenges for both AMC and its organic and industrial suppliers.***

The simulation examined the impact of shifting the burden of the safety stock, holding costs, and pipeline costs into the supply chain. The research showed that investing money in strategically located companies, upstream in the supply chain, and holding more than the minimum stock could lead to shortened service times with an increased capability to meet demand surges and weapon systems readiness requirements.

In the process of developing the new system, many such complexities were revealed and led to interesting research projects by the project team. A recently completed graduate thesis, for instance, investigates the inventory stock-out issue of class IX items (Neidert, 2011). Items with stock-out issues are identified and characterized by statistical analysis. The study proposed a framework and tool for inventory management based on factors such as the inventory level and trend of items. The stock position of an item is forecasted with relatively high accuracy, which would assist supply chain managers in their inventory decisions. For instance, using this framework can help to evaluate whether items are likely to be overstocked or understocked. The model has been well received, and the framework is being used as the basis for a more in-depth analysis of supply chain efficiency.

While the project team made considerable technical progress on the development of a more real-time and integrated metrics system to enhance the current system, the team did not address organizational and cultural elements that are typical in any large-scale organization. However, our observation is that in the time since the project, AMC and subordinate commands have included insights from this and other projects to make important progress on improving their metrics systems.

Conclusion

Integrating performance metrics in an end-to-end supply chain requires a conscientious effort to identify, define, and calculate performance metrics from authorized data sources to provide a broad set of lower level metrics. Integrating lower level metrics into a higher level metrics system requires that data sources from nonintegrated systems be integrated into a system that automatically pulls and accumulates detailed data into local and aggregated measures to provide cascading levels of detail that support local and leadership decision making. Data integrity and cleansing are an integral part of the process, as is reaching common agreement—in this case with the Army—on appropriate data sources, algorithms to calculate metrics, and the design of a dashboard that supports leadership decisions and performance evaluation at higher levels, with drill-down capability for lower level decision making.

Focusing on the end customer, such as warfighters, drove the development of measures such as CWT. To support the best and most capable national defense, we must make sure our warfighters receive the supply support they need at the time they need it. Making sure this process works well over time requires a performance measuring and reporting system that directly supports high performance on measures such as CWT.

The Army and UAH partnership on the eSCALE project led to learning and improvement from both partners and resulted in the development of CWT as the primary high-level metric for on-time performance, which was adopted by JSCA (Fletcher, 2011). Additional advances have been made by UAH, with Army support, on development of an integrated metrics system that includes quality and reliability, and how high CWT and lower levels of quality and reliability negatively affect life-cycle costs, as measured by total ownership cost (Peeler, 2003). Those advancements are the subject of a future article (see Burns & Nicholls, 2013).

References

Beamon, B. M. (1998). Supply chain design and analysis: Models and methods. *International Journal of Production Economics, 55*(3), 281–294.

Bititci, U. S., Carrie, A. S., & McDevitt, L. (1997). Integrated performance measurement systems: A development guide. *International Journal of Operations & Production Management, 17*(5), 522–534.

Bititci, U. S., Carrie, A. S., & Turner, T. (2008). Integrated performance measurement systems: Structure and dynamics. In A. D. Neely (Ed.), *Business Performance Measurement: Theory and Practice* (pp. 174–197). Cambridge, U.K.: Cambridge University Press.

Burns, L., & Nicholls, G. (2013). How government acquisition decisions change when moving from price-based to life cycle cost analysis for high value long life products: Transforming the "more is less" paradox to "less is more." Working paper, pp. 1–40.

Contractor Performance Assessment Reporting System. (2016). *Home page.* Retrieved April 21, 2016, from https://www.cpars.gov/

Department of Defense. (2009). *Department of Defense reliability, availability, maintainability, and cost rationale report manual.* Retrieved from http://www.acq.osd.mil/se/docs/DoD-RAM-C-Manual.pdf

Department of Defense. (2012). *Defense manufacturing management guide for program managers.* Retrieved from https://acc.dau.mil/CommunityBrowser.aspx?id=520540&lang=en-US

Fletcher, M. P. (2011, March–April). Joint Supply Chain Architecture. *Army Sustainment, 43*(2), 20-24. Retrieved from http://www.almc.army.mil/alog/issues/MarApr11/joint_supply_chain.html

Francis, D. (2001). Managing people in agile organisations. In A. Gunasekaran (Ed.), *Agile Manufacturing: The 21st. Century Competitive Strategy* (pp. 193–202). New York, NY: Elsevier.

General Accounting Office. (2003). *Best practices: Setting requirements differently could reduce weapon systems' total ownership costs* (GAO-03-57). Retrieved from http://www.gao.gov/cgi-bin/getrpt?GAO-03-57

Government-Industry Data Exchange Program. (2016). *Home page.* Retrieved from http://www.gidep.org/

Gunasekaran, A., Patel, C., & Ronald, E. M. (2004). A framework for supply chain performance measurement. *International Journal of Production Economics, 87*(3), 333–347.

Henderson, R. (1994). Managing innovation in the information age. *Harvard Business Review, 72*(1), 100–107.

Information Technology Association of America. (2008). *Reliability program standard for systems design, development, and manufacturing* (GEIA-STD-0009, pp. 1–72). Arlington, VA: Author.

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard—Measures that drive performance. *Harvard Business Review, 70*(1), 71–79.

Keebler, J. S., Manrodt, K. B., Durtsche, D. A., & Ledyard, D. M. (1999). *Keeping score: Measuring the business value of logistics in the supply chain.* Oak Brook, IL: Council of Logistics Management.

Melnyk, S. A., Calantone, R. J., Luft, J., Stewart, D. M., Zsidisin, G. A., Hanson, J., & Burns, L. (2005). An empirical investigation of the metrics alignment process. *International Journal of Productivity and Performance Management, 54*(5/6), 312–324.

Melnyk, S. A., Stewart, D. M., & Swink, M. (2004). Metrics and performance measurement in operations management: Dealing with the metrics maze. *Journal of Operations Management, 22*(3), 209–217.

Neidert, L. (2011). *Stock-out as it relates to an army supply chain: An analysis of causes and trends.* (Unpublished master's thesis). University of Alabama, Huntsville, AL.

Oliver, N., Delbridge, R., & Lowe, J. (1996). The European auto components industry. *International Journal of Operations & Production Management, 16*(11), 85.

Past Performance Information Retrieval System. (2016). *Home page.* Retrieved April 21, 2016, from https://www.ppirs.gov/

Peeler, D. L., Jr. (2003). DoD's reduction in total ownership cost effort: R-TOC—A viable solution to DoD's funding predicament. *Armed Forces Comptroller, 48*(2), 18–21.

Siegl, M. B. (2008). *Understanding the supply chain operations reference model. Army Logistician, 40*(3). Retrieved from http://www.alu.army.mil/alog/issues/MayJun08/ref_model_supplychain.html

Solis, W. M. (2009). *Defense inventory: Army needs to evaluate impact of recent actions to improve demand forecasts for spare parts* (GAO-09-199). Retrieved from http://www.gao.gov/products/A85085

Supply Chain Council. (2016). *Supply Chain Operations Reference Model Version 11*. Retrieved from http://www.apics.org/sites/apics-supply-chain-council/frameworks/scor

Biographies

Dr. Fan T. Tseng is professor of Management Science and currently the department chair of Information Systems, Management, and Marketing in the College of Business Administration at the University of Alabama in Huntsville. His current research interests include supply chain analytics and healthcare analytics. He has conducted research in operations research and operations management, and has published in areas including supply chain, healthcare, scheduling, optimization, heuristics, and simulation.

(E-mail address: fan.tseng@uah.edu)

Dr. Laird Burns is associate professor of Management Science and Supply Chain Management in the College of Business at the University of Alabama in Huntsville. He was a group hire between the colleges of business and engineering. His research interests include supply chain design and related work in systems engineering, risk and resilience in supply networks, dynamics of integrated supply networks, supply chain management, quality and reliability, life-cycle costs, agent-based modeling, operations management, and complex systems.

(E-mail address: laird.burns@uah.edu)

Dr. James T. Simpson is distinguished professor emeritus and principal research scientist in the College of Business Administration at the University of Alabama in Huntsville. His research has appeared in leading academic journals such as *Journal of Marketing Research*, *Journal of Business Research*, *Journal of Systems and Software*, *Decision Support Systems*, and *European Journal of Innovation Management*. He has lectured at universities in Russia, China, Romania, Ireland, France, and England.

(E-mail address: james.simpson@uah.edu)

Dr. David Berkowitz is dean of graduate studies and professor of Marketing at the University of Alabama in Huntsville. His research focuses on innovation, product development and supply chain for complex, long life-cycle products. His research has appeared in journals such as the *Journal of Retailing*, *Journal of Advertising Research*, *Journal of Advertising*, *European Journal of Marketing*, *Defense Acquisition Review Journal*, and the *Journal of Product Innovation Management*.

(E-mail address: david.berkowitz@uah.edu)

SCANDAL and TRAGEDY? Or Acquisition Lessons Relearned by the F-35 PROGRAM

Col Roger Witek, USAF (Ret.)

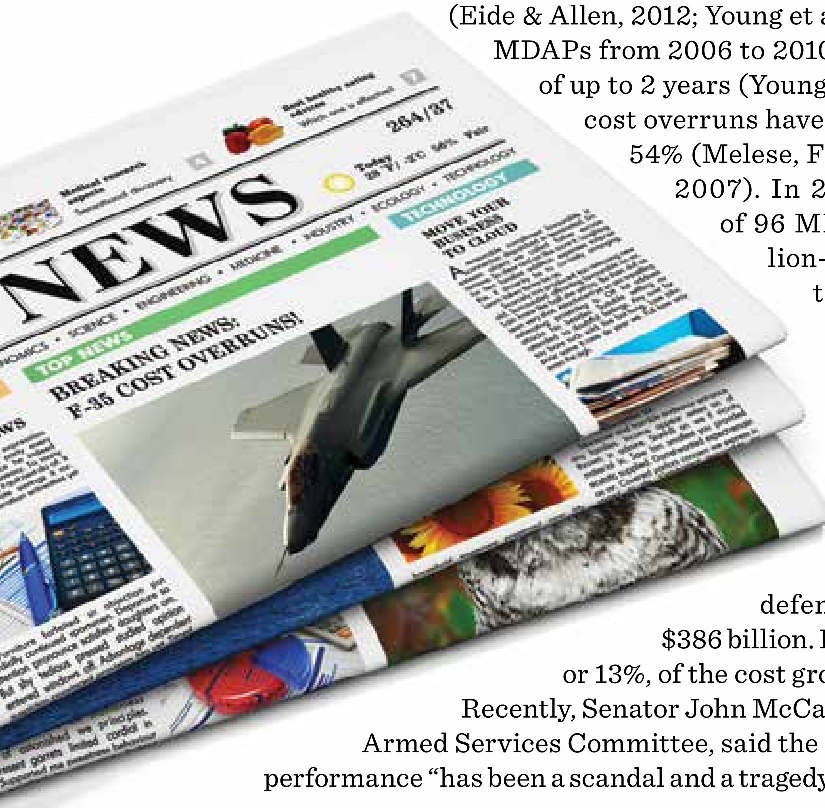
Major defense acquisition programs historically have had difficulty controlling cost, maintaining schedule, and attaining *performance* due to various acquisition strategy challenges. Likewise, with previous joint aircraft programs (F-111, V-22, T-6) and now with the F-35 program, challenges associated with *Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency,* and *Encouraging Partnering* have affected schedule, cost, and performance outcomes. This article summarizes the triangulated research analysis on the comparison of previous joint aircraft acquisition programs, the mining and coding of government agency/think tank reports and scholarly journals on the F-35 program, and the mining and coding of questionnaires given to subject matter experts working on the F-35 program. It argues that the F-35 program has relearned some old lessons and learned some new ones, and it makes recommendations on joint aircraft acquisition strategies for the future to avoid the perception of scandal and tragedy.

DOI: *http://dx.doi.org/10.22594/dau.16-746.24.01*
Keywords: *requirements, technology, commonality, concurrency, partnering*

Image designed by Diane Fleischer

The U.S. Congress has long been concerned about controlling schedule and cost overruns, and attaining expected performance in major defense acquisition programs (MDAP; Blickstein, Nemfakos, & Sollinger, 2013). Schedule, cost, and performance are the three inextricably linked pillars of acquisition (Defense Acquisition University [DAU], 2006).

The 1986 President's Blue Ribbon Commission on Defense Management, referred to as the Packard Commission Report, determined that MDAPs take too long to develop, cost too much, and incorporate obsolete technology by the time they are fielded. More recently, the 2009 Weapon Systems Acquisition Reform Act sought to improve the likelihood of success of MDAPs by focusing on decisions and types of strategies at their inception (Eide & Allen, 2012; Young et al., 2010). Out of 95 known MDAPs from 2006 to 2010, 40% experienced delays of up to 2 years (Young et al., 2010). Total MDAP cost overruns have averaged between 20 and 54% (Melese, Franck, Angelis, & Dillard, 2007). In 2011, the DoD's portfolio of 96 MDAPs stood at $1.58 trillion—$74.4 billion or 5% more than in 2010 (Government Accountability Office [GAO], 2012a). The F-35 Joint Strike Fighter (JSF) program, which seeks to develop and procure 2,457 aircraft for the United States, is the largest single global defense program in history at $386 billion. It accounted for $10 billion, or 13%, of the cost growth in 2011 (GAO, 2011b). Recently, Senator John McCain, chairman of the Senate Armed Services Committee, said the F-35 program's record and performance "has been a scandal and a tragedy" (Associated Press, 2016).



Like previous joint aircraft programs (F-111, V-22, and T-6), the F-35 JSF program has been subject to schedule, cost, and performance shortcomings due to acquisition strategy challenges identified by U.S. Government agencies and U.S. Government think tanks (Blickstein et al., 2013; GAO, 2012c). These acquisition strategy challenges include Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency,

and Encouraging Partnering (Blickstein et al., 2013; Dunne, 2011; Ergas, 2009; GAO, 2012b; Wicht & Crawley, 2012). This article summarizes a 2015 dissertation's research methodology and lessons learned about strategy challenges from a triangulated, qualitative case study analysis on previous joint aircraft acquisition programs, on governmental agency and think tank reports and scholarly journals on the JSF, and on questionnaire responses by subject matter experts (SME) who were currently working on the JSF program. Secretary of Defense (SecDef) Robert Gates (2009) complained that acquisition problems have been persistent and difficult despite congressional involvement in trying to resolve them (more than 125 studies since World War II have produced no comprehensive, effective, and permanent solutions). The goal of this unique research is to improve the cost, schedule, and performance of the JSF program and other MDAPs by understanding better, and making recommendations on, the acquisition strategy challenges of Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency, and Encouraging Partnering. The research shows the F-35 program relearned some old lessons and learned some new ones.

The F-35 is a stealthy, supersonic, multirole fighter built by Lockheed Martin (LM) in three variants to penetrate modern integrated air defenses. The U.S. Air Force (USAF), several Partners, and Foreign Military Sales (FMS) countries will fly the F-35A Conventional Take-Off and Landing (CTOL). The U.S. Marine Corps (USMC), a couple of Partners, and a possible FMS country will fly the F-35B Short Take-Off/Vertical Landing (STOVL). The STOVL is the same size as the CTOL, but the STOVL carries less fuel because of the vertical lift fan. The U.S. Navy (USN) and the USMC will fly the F-35C Carrier Variant with a much larger wing than the other models for carrier landing approach speeds, but it is not being internationally marketed.

Methodology

A triangulated, multilayered, qualitative case study was used to synthesize lessons from three lanes of analysis: the comparison of previous joint aircraft acquisition programs (F-111, V-22, and T-6), the mining and coding of government agency and think tank reports and scholarly journals on the F-35 program, and the mining and coding of questionnaire responses from SMEs who were currently working on the F-35 program at the time of the research in 2015. The main research question of this study was to understand how acquisition strategy challenges (Harnessing Technology, Demanding Commonality, Evoking Concurrency, and Encouraging

Partnering) have both helped and hindered joint aircraft programs' schedule, cost, and performance (the pillars) in terms of the acquisition strategy's original intent, its positive and negative effects on the pillars, and what improvements could be made.

Jogulu and Pansiri (2011) supported triangulation over a single approach because it strengthens findings and inferences made for understanding program management discipline. Yin (1994) and Patton (2002) believed a deep-rooted and multilevel case study analysis could help formulate appropriate relationships between phenomena. As for mining and coding, Patton (2002) offered how to use qualitative data analysis (QDA) software to categorize and to make sense out of massive amounts of data. Finally, submitting a questionnaire with phenomenological attributes to a heuristic group of SMEs produced high-fidelity qualitative analysis of the experiences, beliefs, and perceptions of respondents from multiple perspectives (Moustakas, 1990; Shank, 2006). The questionnaire is phenomenological because the respondents answer the questions as they perceive the situation (Moustakas, 1990). This approach is heuristic because the research process involved the experiences of the researcher in relation to the questionnaire respondents (Moustakas, 1990). As chief of the JSF Coordination and Training Office (JCTO) at the USAF's Air Education and Training Command (AETC) for 4 years, the researcher constantly reflected on and interpreted daily interactions with SMEs from the F-35 Joint Program Office (JPO), LM, and the USAF's headquarters, test community, and training operators. Wacker (1998) concluded that, when dealing with the social sciences (including program management), good recommendations come from open questions, often applying the researcher's and SME's own experiences, instead of from scientific, quantifiable analysis.

Each analysis lane is considered multilayered because three previous joint aircraft acquisition programs were reviewed and because three types of documents and three categories of questionnaires were mined and coded. The F-111, V-22, and T-6 programs were the most suitable comparisons in the number of Services initially interested in combining development, the number of aircraft being procured, and the overall complexity. Potential sampling of previous joint aircraft acquisition programs could have included some well-known and successful aircraft acquisition programs like the F-4, F-5, and A-7 that were used by the Air Force, Navy, and Marines, and several other countries, but they were originally developed by a single Service first (Antill & Ito, 2013; Pike, 2011). Although scholarly journals and think tank reports exist on the F-35 program, General Accounting Office/GAO and Director of Operational Test and Evaluation (DOT&E) reports on the F-35

outnumbered the scholarly journals and think tank reports by 44 to 24 in this study. General Accounting Office/GAO and DOT&E reports on the JSF go back to 2003, when some of the acquisition strategy challenges began to emerge and provided vast amounts of numerical and statistical data, as well as well-defined problems and recommendations. The General Accounting Office/GAO and DOT&E reports on the JSF were prepared by experts in the field of acquisition; since the reports were a matter of public record, this promoted validity and dependability. The original plan was to e-mail the questionnaire to about 50 JSF SMEs in three categories—20 out of 200 from the F-35 JPO as the managers, 10 out of 100 from LM as the providers, and 20 out of 200 from the U.S. military services as the customers. Leedy and Ormrod (2009) recommended an unstructured survey to sample between five and 25 individuals. The 2015 study accepted the upper limit of 25 within the manager and customer categories.

The disclosure that the researcher was AETC's chief of the JCTO during the time of the study promoted credibility and integrity with the questionnaire participants, besides contributing to a heuristic research approach.

As for construct validity on the unstructured questionnaire, the 2015 study followed the approach from a questionnaire in a dissertation by Uda (2012). The questionnaire in the 2015 study was championed by the JPO deputy program executive officer and vetted by the JPO security officer and lawyers. A limitation of the 2015 study was that only unclassified information was used from open source literature and unclassified answers from the questionnaire. One of the toughest challenges to internal validity was the need to guard against the researcher's and respondents' expectancies and biases while being so intimately involved with the F-35 program. The participants were not led to foregone conclusions through an interview; this is why an open questionnaire was used instead. The 2015 study was delimited in scope by using a qualitative methodology instead of a quantitative one, because the large amount of government data would have been difficult to quantify any actionable recommendations. DoD instructions and university institutional review board processes ensured ethical standards were maintained with respect to questionnaire participants to the point where each participant's commander or supervisor approved contact by the researcher.

Previous Joint Aircraft Acquisition Programs

Three previous joint aircraft programs similar in scope to the JSF were compared: the 1960s Tactical Fighter Experimental (TFX) F-111, the 1980s Joint Service Vertical Takeoff & Landing Experimental (JVX) V-22, and the 1990s Joint Primary Aircraft Training System (JPATS) T-6 programs. Examining the effect on the pillars of acquisition by the previous programs' acquisition strategy challenges acted as a precursor to what the JSF program has experienced. As previously discussed, although there were successful aircraft acquisition programs like the F-4, F-5, and A-7, they were developed by a single Service (Antill & Ito, 2013; Pike, 2011). There are also some successful joint missile programs like the Joint Direct Attack Munition, but it simply did not match the scope, scale, and complexities of the F-111, V-22, and T-6 programs.

Tactical Fighter Experimental (TFX) F-111 Program

The TFX program introduced the multimission concept that would affect the attack aircraft industry for the next few decades (Miller, 1982). After World War II, attack aircraft were developed for single purpose missions: nuclear strategic bombing, tactical interdiction, air superiority, or close air support (Miller, 1982). Furthermore, SecDef McNamara wanted to shift doctrine from massive nuclear retaliation to a range of

conventional options (Coulam, 1977). Still, the USAF wanted a follow-on F-105 fighter-bomber for the delivery of internally carried tactical nuclear missiles (*TFX Contract Investigation*, 1963). The USN wanted an air-to-air missile carrier to identify and shoot down enemy planes at extended ranges from their carriers (Coulam, 1977). The USAF and the USN could agree only on a swing-wing, two-seat, and twin-engine design (Art, 1969). The USAF wanted a tandem-seat aircraft (pilot in front and weapon system operator behind) for low-level penetration ground-attack, while the USN wanted a shorter, high-altitude interceptor with side-by-side seating to allow the pilot and radar intercept officer to share the radar display (Miller, 1982). Coulam (1977) concluded that directly competing requirements were inevitably traded off, never fully meeting either Service's requirements.

The TFX program resulted in the F-111, produced by General Dynamics (GD), serving primarily as a supersonic, medium-range interdictor and tactical attack aircraft that later filled the roles of strategic bomber and electronic-warfare aircraft (Logan, 1998). It first entered service with the USAF in 1967 and then with the Royal Australian Air Force in 1973 (Logan, 1998). The F-111 featured new variable-geometry wings for high- and low-speed flight with leading-edge slats and double-slotted flaps over its full length to create more lift for relatively short runway use. It also had afterburning turbofan engines and automated terrain-following radar for low-level, high-speed flight (Logan, 1998). A major failing of the TFX program was that it asked too much of technology too soon (*TFX Contract Investigation*, 1970). SecDef McNamara and the Services looked to GD to solve all of the issues with new innovations (*TFX Contract Investigation*, 1963). Although variable-geometry wings worked as advertised, the poor performance of the Mark II low-level avionics, T-30 turbofan engines, and variable inlets would

plague the military services (*TFX Contract Investigation*, 1970). The program also experienced a 25% concurrency rate—141 out of 547 total USAF F-111s needed retrofits (Coulam, 1977; Richey, 2005).

> ***Most USAF programs in the 1950s exceeded their costs by 100 to 200% and their schedules by 36 to 50% (Summers, 1965).***

Cost estimates also concerned SecDef McNamara (*TFX Contract Investigation*, 1970). He recognized that the Services had limited resources and funding from Congress, so in order to get more from their budgets, they would encourage bids that were unrealistically low (*TFX Contract Investigation*, 1963). Once the Services had congressional support and dedication to continue the program, costs predictably rose, but the Services were likely to get additional funds to finish the program (*TFX Contract Investigation*, 1963). Most USAF programs in the 1950s exceeded their costs by 100 to 200% and their schedules by 36 to 50% (Summers, 1965). McNamara passionately drove for a single aircraft to meet the needs of both the USAF and USN, expecting a high degree of commonality between the two versions (*TFX Contract Investigation*, 1963). GD planned to reduce cost and risk by adding Grumman and P&W as aircraft development partners (*TFX Contract Investigation*, 1963). GD was also able to reduce predicted unit price by courting the U.K. early as a partner to buy several of a specialized variant of the F-111, expecting commonality to save money (Hunter, 1998; Logan, 1998). SecDef McNamara chose GD over Boeing for more realistic cost estimates, but the Services were guilty of assuming high expectations on technical performance (*TFX Contract Investigation*, 1970). In the end, the TFX program had a 100% cost overrun and a 30 to 40% schedule overrun (*TFX Contract Investigation*, 1970), and only seven F-111Bs were built for the USN for test purposes before they cancelled out of the program (General Accounting Office, 1973).

Joint Service Vertical Takeoff and Landing Experimental (JVX) V-22 Program

The V-22 is a multimission, tilt-rotor aircraft with both a vertical takeoff and landing (VTOL) capability like a helicopter and a fixed-wing aircraft capability, achieved by tilting its wing-mounted rotors to act as propellers (General Accounting Office, 1990, 1994; Whittle, 2010). The JVX program

started in 1981 to meet joint Service requirements that would satisfy USMC medium-lift assault, USN search and rescue, and USAF long-range special operations (General Accounting Office, 1990, 1994, 1997). The DoD awarded Bell Helicopter and Boeing Helicopters a development contract in 1983 (Whittle, 2010). The U.S. Army (USA) planned to use the USMC's assault requirements for its medium cargo lift and medical evacuation needs (General Accounting Office, 1986). When the first V-22 rolled out in 1988, the USA had already left the program for good to focus its budget on more immediate aviation programs (Whittle, 2010).

The JVX's original program cost estimates changed significantly, and its development process was long and controversial (Whittle, 2010). One of the USN's cost-saving strategies for the USMC's MV-22s included a high level of concurrent development (General Accounting Office, 1994). The General Accounting Office (1994) warned that such concurrency involved high risk that eventually required rescheduling and spending on increased overtime. The V-22 began flight testing in 1989 and started design alterations immediately (Whittle, 2010). The complexity and difficulties of being the first tilt-rotor intended for military service in the world led to many years of development (Whittle, 2010). The JVX program faced opposition in the Senate in 1989, surviving two separate motions that both could have resulted in program cancellation (Whittle, 2010). The full-scale development contract was even terminated once in October 1992 because Bell and Boeing failed to assemble all six flight-test aircraft, failed to perform all planned drop and fatigue tests, and did not complete all flight testing (General Accounting Office, 1994).

The V-22 program entered full-rate production without mature manufacturing processes that required a redesign and retrofit of the hydraulic and electric system and led to a Nunn-McCurdy breach in 2001 (General Accounting Office, 2003). Although the USMC began crew training for the MV-22 Osprey in 2000, it did not declare initial operational capability (IOC) until 2007 (Whittle, 2010). The Osprey's other current operator, the USAF, declared IOC in 2009 with their CV-22 version of the tilt-rotor (Whittle, 2010). Although 12 MV-22s deployed to Iraq in January 2009, and confirmed there that the MV-22's enhanced speed and range enabled personnel and internally carried cargo to be transported faster and farther than by legacy helicopters (GAO, 2009), almost 30 years had passed from program inception to real-world execution. During that time, V-22 costs have risen sharply above initial projections—1986 estimates (stated in fiscal year 2009 dollars) that the program would build nearly 1,000 aircraft in 10 years at $37.7 million each have shifted to fewer than 500 aircraft at $93.4

million each—a procurement unit cost increase of 148%, while research, development, testing, and evaluation costs increased over 200% (Gertler, 2011). Even after the Department of State approved Japan in 2015 for acquisition of up to 17 V-22B Block-C Ospreys and all the logistical support, Japan deferred their purchase indefinitely due to their budget restraints and the predicted increased costs to maintain the complicated weapon system (McCullough, 2015).

Joint Primary Aircraft Training System (JPATS) T-6 Program

In 1988, the USAF and the USN worked together on the DoD Trainer Aircraft Masterplan and formed the JPATS program to modernize their training aircraft fleets and methods of primary flight training (AETC, 2010). Once the USAF and USN finally agreed on tandem cockpits (the instructor behind the student) and the anthropometrics of the ejection seat to allow more women into flight training, the USAF and USN settled on the commercial-off-the-shelf (COTS) Hawker Beechcraft (formerly Raytheon Aircraft Company) Pilatus PC-9 aircraft (AETC, 2010).

Military-unique design requirements off the COTS baseline grew from about 5% when the program entered limited production in 1995 to almost 70% by the early 2000s (General Accounting Office, 2003). Furthermore,

balancing requirements between the USAF and USN led to a 22% heavier aircraft than its original COTS version (Gantt, 2002). It took 7 years from the establishment of the DoD Trainer Aircraft Masterplan and JPATS program in 1988 to aircraft coming off the assembly line in 1995 for a relatively simple mission (AETC, 2010). Twelve years into production in 2007, JPATS experienced a "critical" Nunn-McCurdy breach, exceeding 50% cost growth from its baseline (GAO, 2007). A DoD review concluded that the cost growth was attributed to changes in government requirements, and the remainder was due to immature and unchecked manufacturing processes (GAO, 2007). Once JPATS was rebaselined in 2008 for cost and schedule as required by Nunn-McCurdy, several foreign countries signed contracts directly with Hawker Beechcraft via direct commercial sales vice FMS (DoD, 2012).

Assessment of Previous Joint Aircraft Programs

The TFX F-111, JVX V-22, and JPATS T-6 programs experienced similar instances of acquisition strategy challenges in relation to the pillars of acquisition. Several specific instances were tabularized in the dissertation study this article is based upon. Each program had at least one acquisition strategy challenge affecting more than one acquisition pillar at the same time or one of the challenges worked in tandem with another to affect either one or more of the pillars. Figure 1 begins to answer the research subquestions concerning acquisition strategy challenges' positive and negative effects on the pillars of acquisition just by looking at the up or down arrows beside each factor below each pillar column in relation to the acquisition strategy challenge on the left. For example, at the top of Figure 1 at the intersection of "Cost" and "Balancing Requirements," a down arrow representing reduced cost stands next to the comment "Combine programs to get one aircraft," while an up arrow stands next to the comment "Meet multi-Services' missions and operating environments." The important message that Figure 1 is trying to convey is that acquisition strategy challenges are not mutually exclusive, exemplified by having the color of one strategy affecting one or more pillars of another strategy. For example, Balancing Requirements (red) and Demanding Commonality (yellow) significantly overlap and affect each other. Cost and performance are highlighted by the opposite color of the strategy challenge because it is difficult to meet the Services' individual mission needs and preferences. Likewise, Harnessing Technology (blue) and Evoking Concurrency (green) overlap, affecting mostly performance and schedule because there can never be enough engineering and modeling of new technology when that new technology is already in the production line with few actual test flights.

FIGURE 1. PREVIOUS JOINT AIRCRAFT PROGRAMS' INTERRELATIONSHIPS BETWEEN PILLARS OF ACQUISITION AND ACQUISITION STRATEGY CHALLENGES

Acquisition Strategy Challenges	Pillars of Acquisition		
	Schedule ▲ = Lengthens ▼ = Shortens	**Cost** ▲ = Increases ▼ = Decreases	**Technical Performance** ▲ = Keeps or Improves ▼ = Reduces
Balancing Requirements	▲ Drive consensus ▲ Rushing to meet milestones	▼ Combine programs to get one aircraft ▲ Meet multi-Services' missions and operating environments ▲ Mission/requirements additions or changes ▲ Contradictory aerodynamic requirements	▼ Contradictory aerodynamic requirements ▼ Concessions eventually made by one Service
Harnessing Technology	▲ Immature manufacturing processes ▲ Not previously attempted ▲ Immature technology ▲ Mature technology or commercial-off-the-shelf	▲ Not previously attempted ▲ Immature technology	▲ Need for asymmetric advantage ▼ Not previously attempted ▼ Immature technology ▲ Mature technology or commercial-off-the-shelf
Demanding Commonality	▲ More stakeholders require more time to decide	▼ One aircraft to meet needs of two or more Services ▲ % of common parts decreases as time passes	▼ As Service autonomy is threatened
Evoking Concurrency	▼ Development, production, testing at nearly the same time ▲ Immature manufacturing processes	▼ Development, production, testing at nearly the same time ▲ Retrofits if problems arise during testing ▲ Immature technology ▲ Mature technology or commercial-off-the-shelf	▼ If not enough testing or the right type of testing—simulated or live flight ▼ Retrofits if problems arise during testing
Encouraging International Partnering	▲ More countries sooner have more say	▼ More countries buying ▼ More countries sooner	▼ Desire for coalition interoperability

Note. Specific instances of strategy effect on pillars from TFX F-111, JVX V-22, and JPATS T-6 programs were generalized from the dissertation. Color scheme demonstrates that acquisition strategy challenges are not mutually exclusive.

Government Agency and Think Tank Reports and Scholarly Journals on the JSF

Government agency and think tank reports and scholarly journals on the JSF program were mined and coded for acquisition strategy challenges and pillars of acquisition, and to see whether and how recommendations were executed. Although ATLAS.ti QDA software assisted in mining and coding, it was mostly used for cataloging and for determining strong associations between acquisition strategy challenges and pillars of acquisition, and within one another by counting co-occurrences and determining c-coefficients (Friese, 2013). Friese (2013) recommended that strong associations should be for c-coefficients ≥ 0.08. The 2015 study addressed each strong association as long as there were two co-occurrences.

> ***Government agency and think tank reports and scholarly journals on the JSF program were mined and coded for acquisition strategy challenges and pillars of acquisition, and to see whether and how recommendations were executed.***

Within the "super-family" code of *acquisition strategy challenges*, the following coding families were established: Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency, and Encouraging Partnering. Although the research questionnaire asked three questions on each strategy, there were really four coding components for each family code: "original intent, negative effects, positive effects, and improvement recommendations." Within ATLAS.ti, they were coded as Balancing Requirements Intent (BRI), Balancing Requirements Negative (BRN), Balancing Requirements Positive (BRP), Balancing Requirements Recommendation (BRR), Technology Intent (TI), Technology Negative (TN), Technology Positive (TP), Technology Recommendation (TR), Commonality Intent (CMI), Commonality Negative (CMN), Commonality Positive (CMP), Commonality Recommendation (CMR), Concurrency Intent (CCI), Concurrency Negative (CCN), Concurrency Positive (CCP), Concurrency Recommendation (CCR), Partnering Intent (PTI), Partnering Negative (PTN), Partnering Positive (PTP), and Partnering Recommendation (PTR).

Within the "super-family" *pillars of acquisition code*, the coding families of "cost, schedule, and performance" were created. After the literature review and pre-reading the respondents' questionnaires, it was appropriate to have five coding components for each pillar family code to separate the reasons from other statements. For cost, the codes were Cost Increased (CI)—bad, Cost Increase Reason (CIR)—why, Cost Decreased (CD)—good, Cost Decrease Reason (CDR)—why, and Cost Recommendation (CR). For schedule, the codes were Schedule Lengthened (SL)—bad, Schedule Lengthened Reason (SLR)—why, Schedule Shortened (SS)—good, Schedule Shortened Reason (SSR)—why, and Schedule Recommendation (SR). For performance, the codes were Performance Reduced (PRD)—bad, Performance Reduced Reason (PRR)—why, Performance Improved (PI)—good, Performance Improved Reason (PIR)—why, and Performance Recommendation (PR).

Table 1 shows the breakdown of family codes by the type of government report or scholarly journal. As previously mentioned, Government Accountability Office (GAO) and DOT&E reports on the F-35 outnumbered the scholarly journals and think tank reports by 44 to 24 in this study. Out of 68 total documents, the GAO was the largest represented group, with 33 total documents between *Selected Acquisition Reports* and specific reports on the F-35. Although the GAO had more documents than the scholarly journals, the GAO usually had the same researchers investigating and writing the reports for several years, so there is a reputation of expertise that could not be ignored. With only 11 documents, DOT&E was coded the most, and its highest percentage of codes went to the Harnessing Technology family of codes. DOT&E had numerous Technology Recommendation (TR) codes to adjudicate the large number of Technology Negative (TN) codes. The GAO also coded the Harnessing Technology family of codes pretty often, but it highlighted mostly TN aspects, some Technology Positive (TP) aspects, and very few Technology Recommendations (TR) as compared to DOT&E. Think tank coding and scholarly journal coding were evenly distributed among the strategies and pillars, but scholarly journals coded the Encouraging Partnering family of codes the most. It did not go unnoticed that the government agency/think tank reports coded more negatively across most strategies; whereas scholarly journals only slightly favored negative codes, but in terms of broad program management. There were 234 passages coded as recommendations from the governmental agency/think tank reports and scholarly journals out of 798 total codes.

TABLE 1. REPORTS'/JOURNALS' FAMILY CODE BREAKDOWN

	DOT&E (11)	GAO *SAR* (13)	GAO F-35 (20)	Think Tanks (8)	Journals (16)	Totals (68)
Balancing Requirements Intent (BRI)	5	0	1	4	0	10
Balancing Requirements Negative (BRN)	12	4	0	7	2	25
Balancing Requirements Positive (BRP)	8	1	0	0	1	10
Balancing Requirements Recommendation (BRR)	6	0	2	4	1	13
Balancing Requirements Sub-total	30	5	3	14	4	56
Technology Intent (TI)	8	2	0	3	1	14
Technology Negative (TN)	39	38	19	10	8	114
Technology Positive (TP)	16	15	3	0	2	36
Technology Recommendation (TR)	61	1	8	7	5	82
Harnessing Technology Sub-total	112	55	30	20	15	232
Commonality Intent (CMI)	0	2	0	5	4	11
Commonality Negative (CMN)	3	0	1	11	7	22
Commonality Positive (CMP)	1	1	0	0	7	9
Commonality Recommendation (CMR)	1	0	1	3	5	10
Demanding Commonality Sub-total	5	3	2	17	21	48
Concurrency Intent (CCI)	0	3	0	2	1	6
Concurrency Negative (CCN)	13	17	15	2	3	50
Concurrency Positive (CCP)	0	3	1	0	0	4
Concurrency Recommendation (CCR)	11	2	6	0	4	23
Evoking Concurrency Sub-total	24	23	22	4	8	81
Partnering Intent (PTI)	0	0	0	3	6	9
Partnering Negative (PTN)	0	1	0	4	8	13
Partnering Positive (PTP)	0	0	0	1	9	10
Partnering Recommendation (PTR)	2	1	0	2	0	5
Encouraging Partnering Sub-total	2	1	0	8	19	30

TABLE 1. REPORTS'/JOURNALS' FAMILY CODE BREAKDOWN, CONTINUED

	DOT&E (11)	GAO *SAR* (13)	GAO F-35 (20)	Think Tanks (8)	Journals (16)	Totals (68)
Cost Increase (CI)—bad	0	8	10	6	5	29
Cost Increase Reason (CIR)—why	4	16	8	8	10	46
Cost Decrease (CD)—good	0	2	1	0	1	4
Cost Decrease Reason (CDR)—why	0	3	3	2	5	13
Cost Recommendation (CR)	9	3	15	13	7	37
Cost Sub-total	12	30	36	19	28	126
Schedule Lengthened (SL)—bad	7	12	9	2	2	32
Schedule Lengthened Reason (SLR)—why	13	20	8	2	2	45
Schedule Shortened (SS)—good	0	1	1	0	0	2
Schedule Shortened Reason (SSR)—why	3	3	2	0	0	8
Schedule Recommendation (SR)	21	3	5	1	2	32
Schedule Sub-total	42	36	25	5	6	115
Performance Reduced (PRD)—bad	16	4	3	4	2	29
Performance Reduced Reason (PRR)—why	10	7	0	1	4	22
Performance Improved (PI)—good	8	8	8	0	1	26
Performance Increased Reason (PIR)—why	2	3	2	0	2	9
Performance Recommendation (PR)	20	1	3	5	3	32
Performance Sub-total	52	21	17	10	12	112
Grand Totals	279	174	135	97	113	798

Note. (X) = number of documents, DOT&E = Director of Operational Test and Evaluation, GAO = Government Accountability Office, *SAR* = *Selected Acquisition Reports*

Table 2 provides insight into the opening (O), working (W), and closing (C) timing of 80 official recommendations made by the DOT&E, the GAO, and the DoD Inspector General directly to the JSF JPO. Most of DOT&E's closed recommendations to the JPO were divided between improvements for flight test strategy, planning and realism, and with Autonomic Logistics Information System (ALIS) testing. DOT&E's open recommendations deal with mission data load test integration, the helmet mounted display system, fueldraulics' survivability, VSim verification, aircraft repair times, Block 2B weapons delivery accuracy, and F-35B STOVL fielding concerns. Most of GAO's closed recommendations to the JPO dealt with cost and reschedule baselining, maintaining expected funding, monitoring software, and F-35B progress. GAO's open recommendations deal with executing knowledge-based, evolutionary acquisitions and limiting production strategies. The open recommendations by the Department of Defense Inspector General (DoD IG) to the JPO deal with establishing quality assurance programs over the contractor, which the JPO disagrees with because it does not have the resources or the responsibility to perform this through the supply chain (DoD IG, 2013).

TABLE 2. TRACKING RECOMMENDATIONS TO THE JPO

Recommendations	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
DOT&E 05—Have data collection and mission replay to evaluate mission effectiveness	O	C									
DOT&E 05—Test and Evaluation Master Plan (TEMP) should identify shortfalls in opposing force/threats	O	W	W	W	W	C					
DOT&E 05—Align requirements for each block aircraft	O	C									
DOT&E 05—Develop predictive model on engine performance after "quick dump" fuel ingestion	O	W	W	C							
DOT&E 05—Reduce fuel ingestion vulnerability by improving fuel bladders around inlet ducts	O	C									
DOT&E 06—Update issues from recent operational assessment		O	C								
DOT&E 06—Consider doing Initial Operational Test and Evaluation (IOT&E) earlier with operationally representative aircraft		O	C								
DOT&E 06—Follow May 06 Defense Acquisition Board ideas for Partner testing		O	C								
DOT&E 06—Fund adequate full-scale aerial target to confirm operational effectiveness		O	C								
DOT&E 06—Conduct full-up, live system-level live-fire ballistic tests on F-35 to determine vulnerability		O	W	C							
DOT&E 07—Retain last two system development and demonstration (SDD) aircraft			O	C							
DOT&E 07—Ensure labs are resourced to execute verification strategy and to surge			O	W	W	W	C				
DOT&E 07—Develop metrics for verification strategy			O	W	W	C					
DOT&E 07—Develop entrance criteria for IOT&E			O	W	W	C					
DOT&E 07—Reinstate dry bay engine fire suppression			O	C							
DOT&E 07—Reinstate engine fuel ingestion suppression liner			O	C							
DOT&E 07—Add/Improve cockpit warning lights to F-35B for ballistic damage before vertical landing			O	W	W	W	W	W	W	C	
DOT&E 07—Retain engine fueldraulics and liquid cooling shutoff valves			O	C							
DOT&E 08—Add resources to flight testing in FY09-11				O	W	W	C				
DOT&E 08—Explain all test changes to DOT&E				O	W	C					

TABLE 2. TRACKING RECOMMENDATIONS TO THE JPO, CONTINUED

Recommendations	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
DOT&E 08—Initiate a Test Plan Working Group				O	W	C					
DOT&E 08—Stabilize production and deliveries of systems needed for OT&E and training				O	W	W	C				
DOT&E 08—Complete third iteration of the TEMP				O	W	C					
DOT&E 08—Improve verifications simulator (VSim) to meet adequate verification testing				O	W	W	C				
DOT&E 08—Restore capability to minimize fueldraulics spillage from threat-induced damage				O	W	W	W	W	W	?	
DOT&E 09—Focus delivery efforts on SDD aircraft					O	W	C				
DOT&E 09—By an Operational Test Review Team, review IOT&E test plan for Block-3 aircraft systems					O	C					
DOT&E 09—Have more transparent contract negotiations for Block-3 test aircraft					O	C					
DOT&E 09—Verify, validate, and accredit test labs					O	W	C				
DOT& E 10—Assure new flight test schedule is realistic						O	W	W	W	?	
DOT&E 10—Evaluate flight test schedule, executed versus planned						O	W	?			
DOT&E 10—Determine impact of technical issues of helmet-mounted display						O	W	?			
DOT&E 10—Assure software integration is in flight test						O	C				
DOT&E 10—Verify/validate msn data loads (MDLs)						O	W	?			
DOT&E 10—Redesign On-board Inert Gas Generating System (OBIGGS) to maintain oxygen levels below where fire can be sustained						O	W	W	W	W	—
DOT&E 11—Use event-driven criteria to begin flt ops							O	?			
DOT&E 11—Test transonic buffeting							O	?			
DOT&E 11—Determine impacts of late structural durability testing							O	?			
DOT&E 11—Improve spare/resupply for flight test							O	?			
DOT&E 11—Survey test plans for certifications by outside government agencies							O	?			
DOT&E 12—Make corrections to Version 4 of TEMP								O	?		

TABLE 2. TRACKING RECOMMENDATIONS TO THE JPO, CONTINUED

Recommendations	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
DOT&E 12—Conduct testing on Autonomic Logistics Information System (ALIS)								O	W	?	
DOT&E 12—Make operational test (OT) aircraft fully production-representative								O	C		
DOT&E 12—Ensure contractor meeting VSim requirements								O	W	?	
DOT&E 12—Assure integrated VSim and MDL testing								O	C		
DOT&E 12—Continue PAO shutoff valve redesign								O	W	?	
DOT&E 12—Consider removing fueldraulic fuses								O	C		
DOT&E 12—Consider keeping dry bay fire extinguisher for the Integrated Power Pack only								O	W	?	
DOT&E 12—Determine ballistic event survival time								O	W	?	
DOT&E 13—Account for historical growth of flight test									O	?	
DOT&E 13—Get VSim data for SDD flight test									O	?	
DOT&E 13—Track metrics for software stability									O	?	
DOT&E 13—Determine viability of putting 270-volt power on 28-volt signal bus									O	C	
DOT&E 13—Track low observable (LO) and non-LO repair times									O	C	
DOT&E 14—Update IOT&E Schedules										O	—
DOT&E 14—Complete MDL testing before flight test										O	—
DOT&E 14—Complete Block 2B weapon delivery accuracy										O	—
DOT&E 14—Require contractor to do finite element analysis on F-35B bulkhead										O	—
DOT&E 14—Resource Block 3 VSim adequately										O	—
DOT&E 14—Accelerate joint technical data (JTD) verification for fielded F-35Bs										O	—
DOT&E 14—Extend decontamination tests										O	—
GAO 05—Establish an executable program consistent with best practices and DoD policy regarding knowledge-based, evolutionary acquisition	O	W	W	W	W	W	W	W	W	W	—
GAO 07—Limit annual production quantities to no more than 24 aircraft per year until each variant's basic flying qualities have been demonstrated			O	W	W	W	W	W	W	W	—

TABLE 2. TRACKING RECOMMENDATIONS TO THE JPO, CONTINUED

Recommendations	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
GAO 08—Revisit and revise the Mid-Course Risk Reduction Plan to address concerns about testing, use of management reserves, and manufacturing				O	W	W	W	W	C		
GAO 08—Improve JSF cost estimate reliability				O	W	W	W	W	C		
GAO 09—Report to the congressional defense committees on the risks and mitigation strategy for use of cost reimbursement contracts for procurement and plans to transition to fixed price contracts					O	W	W	W	C		
GAO 09—Ensure contractor performs periodic schedule risk analyses to improve schedule and budget actions					O	W	W	W	C		
GAO 10—Make a new, comprehensive, and independent assessment of the costs and schedule to complete the program, including military construction, JSF-related expenses in other budgets, and life-cycle costs						O	W	W	C		
GAO 10—Reassess warfighter requirements and, if necessary, defer some capabilities to future increments						O	W	W	C		
GAO 11—Maintain future funding at current levels							O	W	C		
GAO 11—Establish criteria for evaluating the F-35B's progress and make independent reviews, allowing each variant to proceed at its own pace							O	W	C		
GAO 11—Conduct an independent review of the software development and lab accreditation processes							O	W	C		
GAO 12–13—Restructure JSF program by incorporating previous recommendations from GAO 2008–11									*		
GAO 14—Assess/identify specific capabilities that can be delivered to the military services to support their respective initial operational capabilities by July 2015										O	—
GAO 14—Assess the affordability of F-35's current procurement plan that reflects various assumptions about technical progress and future funding										O	—
DoD IG 13—Ensure LM's design and material changes are with government concurrence									O	C	
DoD IG 13—Perform process proofing of supply chain's critical processes and requirements flow verification									O	W	W
DoD IG 13—Establish independent quality assurance organization to review supplier processes									O	W	W

TABLE 2. TRACKING RECOMMENDATIONS TO THE JPO, CONTINUED

Recommendations	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
DoD IG 15—Evaluate open variance on F135 engine											O
DoD IG 15—Resolve nonconformities of P&W's engine software quality management systems											O

Note. Derived and summarized from DOT&E (2005–14), USGAO (2005, 2007–14), DoDIG (2013–15), DOT&E = Director of Operational Test and Evaluation, GAO = Government Accountability Office, IG = Inspector General, PAO = Polyalphaolefin, TEMP = Test and Evaluation Master Plan, IOT&E = Initial Operational Test and Evaluation, SDD = system development and demonstration, VSim = verification simulator, ALIS = Autonomic Logistics Information System, OT = operational test, LO = low observable, msn = mission, flt = flight, O = open, W = working, C = Close, ? = unknown, — = not published yet, * = summary

Figure 2 shows three dense reports'/journals' co-occurrence/c-coefficient tables generated by ATLAS.ti QDA software; these were explained in depth by the dissertation this article is based upon. The table on the left shows the associations between acquisition strategy challenges and the pillars of acquisition. The middle table shows the interassociations within acquisition strategy challenges. The table on the right shows the interassociations within pillars of acquisition. For each table, each coded component shows the number of codes in parentheses along the header rows and columns. At each intersection, the co-occurrences and c-coefficients are paired. The intent here is to show some specific examples to eventually discern some obvious general trends. Data that are covered up by overlapping tables are not significant.

FIGURE 2. REPORTS'/JOURNALS' CO-OCCURRENCE/C-COEFFICIENT TABLES

	CI (29)	CIR (46)	CD (4)	CDR (13)	CR (37)	SL (32)	SLR (45)	SS (2)	SSR (8)	SR (32)	PRD (29)	PRR (22)	PI (26)	PIR (9)	PR (32)
BRI (10)	0	0	0	1/.05	2/.04	0	0	0	0	1/.02	0	0	0	0	0
BRN (25)	2/.04	3/.04	0	0	1/.02	1/.02	1/.01	0	0	0	3/.06	0	0	0	0
BRP (10)	0	0	0	0	1/.02	0	0	0	2/.13	1/.02	0	0	1/.03	0	0
BRR (13)	0	0	0	0	0	0	0	0	0	1/.02	1/.02	0	0	0	1/.02
TI (14)	0	0	0	0	0	0	0	0	1/.05	0	0	0	0	0	0
TN (114)	5/.04	10/.07	0	0	1/.01	10.07	27/.2	0	0	2/.01	13/.1	11/.09	1/.01	0	1/.01
TP (36)	1/.02	0	1/.03* (5)	1/.02											
TR (82)	0	0	0	0											
CMI (11)	0	0	0	4/.2											
CMN (22)	3/.06	7/.11	0	0											
CMP (9)	0	0	0	3/.16											
CMR (10)	0	0	0	0											
CCI (6)	0	0	0	0											
CCN (50)	1/.01	7/.08	0	0											
CCP (4)	0	0	0	0											
CCR (23)	1/.02	1/.01	0	0											
PTI (9)	0	0	0	3/.16											
PTN (13)	1/.02	1/.02	0	0											
PTP (10)	01/.03	0	1/.08	1/.05											
PTR (5)	0	0	0	1/.06											

	BRI (10)	BRN (25)	BRP (10)	BRR (13)	TI (14)	TN (114)	TP (36)	TR (82)	CMI (11)	CMN (22)	CMP (9)	CMR (10)
BRI (10)	0	1/.03	0	1/.05	1/.04	0	0	0	4/.24	0	0	1/.05
BRN (25)	-	0	1/.03	0	0	9/.07	0	0	0	5/.12	0	0
BRP (10)	-	-	0	0	0	0	1/.02	0	0	0	3/.19	0
BRR (13)	-	-	-	0	0	0	0	3/.03 *(6)	1/.04	0	0	2/.1
TI (14)	-											
TN (114)	-											
TP (36)	-											
TR (82)	-											
CMI (11)	-											
CMN (22)	-											
CMP (9)	-											
CMR (10)	-											
CCI (6)	-											
CCN (50)	-											

	CI (30)	CIR (46)	CD (4)	CDR (13)	CR (37)	SL (32)	SLR (45)	SS (2)	SSR (8)
CI (30)	0	1/.01	0	0	1/.02	15/.33	0	0	0
CIR (46)	-	0	0	0	2/.02	2/.03	9/.11	0	1/.02
CD (4)	-	-	0	0	0	0	0	1/.25	0
CDR (13)	-	-	-	0	1/.02	0	0	0	1/.05
CR (37)	-	-	-	-	0	1/.01	1/.01	0	0
SL (32)	-	-	-	-	-	0	2/.03	0	0
SLR (45)	-	-	-	-	-	-	0	0	0
SS (2)	-	-	-	-	-	-	-	0	0

Note. The left table shows strategies versus pillars. The middle table shows strategies' interassociations. The right table shows pillars' interassociations. Data that are covered up by overlapping tables are not significant.

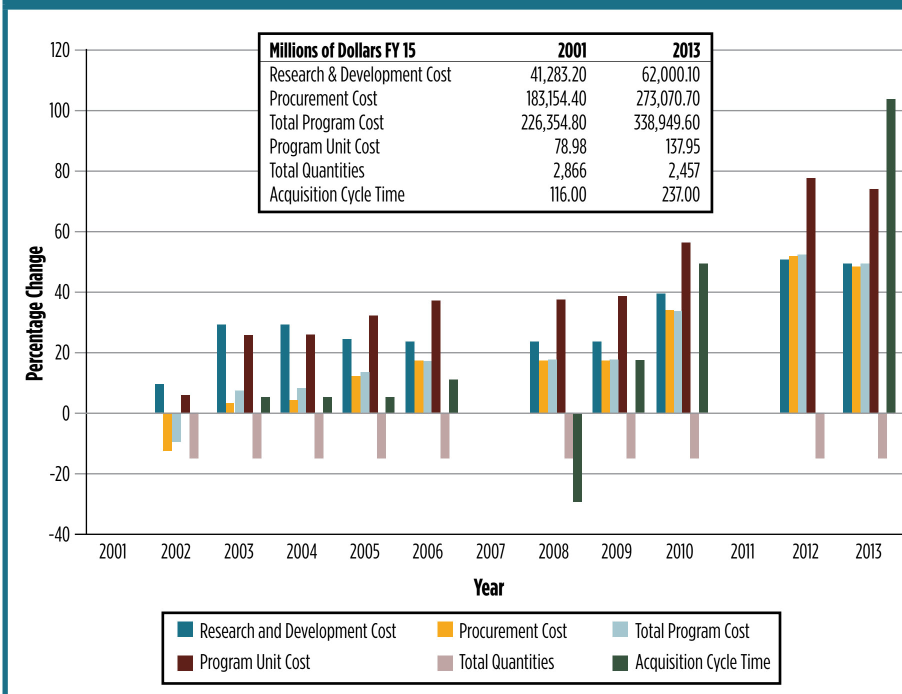
(n), for row/column headers = number of occurrences that were coded, n/c within table where n = co-occurrence and c = c-coefficient—bolded represents strong association if c > .08 and at least two co-occurrences, "*" represents a c-coefficient that may require further investigation—bolded if ratio (r) > 10 and at least two co-occurrences, BRI = Balancing Requirements Intent, BRN = Balancing Requirements Negative, BRP = Balancing Requirements Positive, BRR = Balancing Requirements Recommendation, TI = Technology Intent, TN = Technology Negative, TP = Technology Positive, TR = Technology Recommendation, CMI = Commonality Intent, CMN = Commonality Negative, CMP = Commonality Positive, CMR = Commonality Recommendation, CCI = Concurrency Intent, CCN = Concurrency Negative, CCP = Concurrency Positive, CCR = Concurrency Recommendation, PTI = Partnering Intent, PTN = Partnering Negative, PTP = Partnering Positive, PTR = Partnering Recommendation

The left table in Figure 2 has acquisition strategy challenges' component codes depicted in the left column and the pillars of acquisition component codes depicted in the top row. Strong associations have a c-coefficient ≥ 0.08 and with two co-occurrences. Circled in red, out of the 27 co-occurrences between TN (Technology Negative) and SLR (Schedule Lengthened Reason), eight co-occurrences blamed design immaturity, eight co-occurrences blamed the lack of flight-testing assets that lengthened schedules, six co-occurrences blamed complex software, and three co-occurrences blamed weight control. In the middle table, acquisition strategy challenges interassociations, the one example circled in green, three co-occurrences between BRP (Balancing Requirements Positive) and CMP (Commonality Positive) pointed to Joint Services' expected savings from economies of scale. In the right table, pillars of acquisition interassociations, the one example circled in red, out of the 15 co-occurrences between CI (Cost Increase) and SL (Schedule Lengthened), 12 co-occurrences indicate that schedule delays preceded cost growth, but six out of the 12 pointed to technology challenges being the root cause.

Figure 3 shows percent growth or decline in JSF F-35 reported program cost, unit cost, and acquisition cycle time (official program start to full-rate production) between 2001 and 2013 as reported by the General Accounting Office and GAO. The percentage represented in vertical bars is based on fiscal year (FY) 2015 dollars, but 2001 is the base year shown by 0% change. The 2013 callout box is also in FY 2015 dollars. After the large spike in unit cost in 2003, the JSF program was restructured, but the program triggered a Nunn-McCurdy breach in 2007 for exceeding the 2003 restructured baseline for total program cost and again in 2010 for exceeding the 2007 restructured baseline (GAO, 2014). From 2001 to 2013 in FY 2015 dollars,

the JSF program's total program cost grew 49.7%, its program unit cost grew 74.7%, and the acquisition cycle time increased 104.3%, doubling the time (General Accounting Office 2003, 2004; Government Accountability Office 2005–2011a, 2012a, 2013–2015).

FIGURE 3. JSF PROGRAM PERCENT CHANGE IN COSTS, QUANTITIES, AND SCHEDULE



Note. Adapted from General Accounting Office (2003, 2004) and Government Accountability Office (2005–2011a, 2012a, 2013–2015), FY = fiscal year

Questionnaire Responses from JSF SMEs

The breakdown of questionnaire respondents was slightly different than planned because potential participants were coming and going while the author was coordinating for DoD Information Management Control Office's (IMCO) approval for 4 months. Seventeen questionnaires were sent to the *managers* at the JPO, 25 to the *customers* from the military services, and eight to the *providers*, the contractors at LM. If there had been more than 10 providers (LM contractors), that would have incurred more IMCO requirements because it would have constituted a public survey.

Each participant was allowed 4 weeks beginning in late March of 2015. A 50% response rate was expected to have purposeful sampling according to Zikmund, Babin, Carr, and Griffin (2010) and Creswell (2007). Of the 50 potential participants, 42 participants responded in April 2015 for an incredible 86% response rate—10 out of 17 from the JPO, 25 out of 25 from the Services, and seven out of eight from LM (see Table 3). However, it took 5 weeks for the seven LM responses to be technically reviewed and legally released by the JPO Security Office and lawyers.

TABLE 3. PARTICIPANT RESPONDENT QUESTIONNAIRE TIMING

Participant #	Total Respondents	Group	Day #1	Completed	Day #28	Released
1	1	Manager	19 Mar 15	10 Apr 15	16 Apr 15	N/A
5	2	Manager	19 Mar 15	9 Apr 15	16 Apr 15	N/A
7	3	Manager	19 Mar 15	21 Apr 15*	16 Apr 15	N/A
8	4	Manager	19 Mar 15	13 Apr 15	16 Apr 15	N/A
10	5	Manager	19 Mar 15	14 Apr 15	16 Apr 15	N/A
12	6	Manager	19 Mar 15	13 Apr 15	16 Apr 15	N/A
14	7	Manager	19 Mar 15	13 Apr 15	16 Apr 15	N/A
15	8	Manager	19 Mar 15	14 Apr 15	16 Apr 15	N/A
16	9	Manager	19 Mar 15	15 Apr 15	16 Apr 15	N/A
17	10	Manager	19 Mar 15	16 Apr 15	16 Apr 15	N/A
18	11	Customer	20 Mar 15	10 Apr 15	17 Apr 15	N/A
19	12	Customer	20 Mar 15	17 Apr 15	17 Apr 15	N/A
20	13	Customer	20 Mar 15	17 Apr 15	17 Apr 15	N/A
21	14	Customer	20 Mar 15	16 Apr 16	17 Apr 15	N/A
22	15	Customer	21 Mar 15	20 Apr 15*	18 Apr 15	N/A
23	16	Customer	20 Mar 15	3 Apr 15	17 Apr 15	N/A
24	17	Customer	20 Mar 15	12 Apr 15	17 Apr 15	N/A
25	18	Customer	19 Mar 15	14 Apr 15	16 Apr 15	N/A
26	19	Customer	21 Mar 15	18 Apr 15	18 Apr 15	N/A
27	20	Customer	20 Mar 15	6 Apr 15	17 Apr 15	N/A

TABLE 3. PARTICIPANT RESPONDENT QUESTIONNAIRE TIMING, CONTINUED

Participant #	Total Respondents	Group	Day #1	Completed	Day #28	Released
28	21	Customer	14 Apr 15	27 Apr 15	12 May 15	N/A
29	22	Customer	19 Mar 15	9 Apr 15	16 Apr 15	N/A
30	23	Customer	23 Mar 15	10 Apr 15	20 Apr 15	N/A
31	24	Customer	20 Mar 15	17 Apr 15	17 Apr 15	N/A
32	25	Customer	20 Mar 15	16 Apr 15	17 Apr 15	N/A
33	26	Customer	20 Mar 15	16 Apr 15	17 Apr 15	N/A
34	27	Customer	21 Mar 15	14 Apr 15	18 Apr 15	N/A
35	28	Customer	31 Mar 15	27 Apr 15	28 Apr 15	N/A
36	29	Customer	21 Mar 15	14 Apr 15	18 Apr 15	N/A
37	30	Customer	21 Mar 15	7 Apr 15	18 Apr 15	N/A
38	31	Customer	21 Mar 15	9 Apr 15	18 Apr 15	N/A
39	32	Customer	21 Mar 15	17 Apr 15	18 Apr 15	N/A
40	33	Customer	22 Mar 15	19 Apr 15	19 Apr 15	N/A
41	34	Customer	21 Mar 15	18 Apr 15	18 Apr 15	N/A
42	35	Customer	26 Mar 15	30 Mar 15	23 Apr 15	N/A
43	36	Provider	19 Mar 15	15 Apr 15	16 Apr 15	26 May 15
44	37	Provider	19 Mar 15	15 Apr 15	16 Apr 15	26 May 15
46	38	Provider	19 Mar 15	7 Apr 15	16 Apr 15	26 May 15
47	39	Provider	19 Mar 15	16 Apr 15	16 Apr 15	26 May 15
48	40	Provider	19 Mar 15	8 Apr 15	16 Apr 15	26 May 15
49	41	Provider	19 Mar 15	8 Apr 15	16 Apr 15	26 May 15
50	42	Provider	19 Mar 15	3 Apr 15	16 Apr 15	26 May 15

Note. *Late submission allowed by Dissertation Chair

Table 4 speaks to the credibility of the respondents. From the top of Table 4, although there were only seven LM contractor ("Con") respondents due to not making the questionnaire a public survey, active duty (AD) and general schedule (GS) civilian respondents were practically even at 17 and 18, respectively. A majority (34 of 42 or 81%) of the respondents had earned a master's degree as their highest level of education. Most of the respondents were either executives/directors, program managers (of the variant or subsystem), or action officers with a few dedicated pilots, and a tester. A strong majority (69%) of the respondents considered themselves SMEs, a third of the respondents had pilot experience, a few considered themselves as maintainers (12%), and several (19%) as policy deciders. By far, the providers from LM had the most years of experience in dealing with each acquisition strategy—40% more than the managers at the JPO, and more than double the customers from the Services. Overall, respondents had the most years of experience when dealing with Harnessing Technology, followed by Balancing Requirements, followed by a near three-way tie between the remaining strategy challenges. The large standard deviations in years of experience was because several respondents (12 of 42, 29%) had zero experience with more than one associated acquisition strategy challenge, but they were all highly recommended by their superiors to be invited to participate in the study.

TABLE 4. BREAKDOWN OF RESPONDENTS' DEMOGRAPHICS

Org:	JPO	Mil	LM	Tot	Employee:	AD	GS	Con	Tot	Education:	HS	BS	MS	Dctr	Tot
JPO	10	0	0	10	JPO	5	5	0	10	JPO	0	1	8	1	10
Mil	0	25	0	25	Mil	12	13	0	25	Mil	1	2	21	1	25
LM	0	0	7	7	LM	0	0	7	7	LM	0	1	5	1	7
Total	**10**	**25**	**7**	**42**	**Total**	**17**	**18**	**7**	**42**	**Total**	**1**	**4**	**34**	**3**	**42**

Job (One)	Exec/ Dir	PM	AO	Tester	Op/Trn	Tot	Function (Any)	Pol/Dec	SME	Pilot	Mx
JPO	6	4	0	0	0	10	JPO	30	70	0	0
Mil	0	4	15	1	5	25	Mil	16	64	48	20
LM	5	2	0	0	0	7	LM	29	86	14	0
Total #	**11**	**10**	**15**	**1**	**5**	**42**	**Total %**	**19**	**69**	**33**	**12**

Year's Experience (Avg +/- StD)	Requirements	Technology	Commonality	Concurrency	Partnering	Respondent
JPO (*Manager*)	11.1 +/- 11.4	17.4 +/- 10.9	6.5 +/- 6.4	6.3 +/- 5.3	8.8 +/- 7.1	10.0 +/- 5.8
Mil (*Customer*)	7.8 +/- 7.7	9.5 +/-9.8	4.2 +/- 4.3	5.7 +/- 5.4	4.3 +/- 4.3	6.3 +/- 4.7
LM (*Provider*)	17.0 +/- 9.5	20.0 +/- 14.0	14.1 +/- 9.7	9.1 +/- 5.3	10.9 +/- 10.7	14.2 +/- 7.9
Total	**10.1 +/- 9.4**	**13.2 +/- 11.5**	**6.4 +/- 6.8**	**6.4 +/- 5.4**	**6.4 +/- 6.7**	**8.5 +/- 6.3**

Note. JPO = JSF Program Office, Mil = military services, LM = Lockheed Martin, Tot = total, AD = active duty, GS = general schedule, Con = contractor, HS = high school, BS = bachelor's, MS = master's, Dctr = doctorate, Exec = executive, Dir = director, PM = program manager, AO = action officer, Op = operator, Trn = Trainer, Pol Dec = policy decider, SME = subject matter expert, Mx = maintainer, Avg = average, StD = standard deviation

Similar to "years' experience" from the bottom of Table 4, Table 5 shows the average amount of words respondents used to answer the questions. The customers had the most to say at 1,520 words per an entire set of answers, almost 70% more than the managers, and almost twice as much as the providers. This order held true for the number of words per each acquisition strategy question except for commonality, where the managers and providers switched places for second and third.

Respondents had the most to say about Balancing Requirements, with an average of 114 words for that section of three questions, followed by Harnessing Technology and Encouraging Partnering, with 89 and 83 words per section, respectively, followed by Evoking Concurrency and Demanding Commonality, with 75 and 55 words, respectively. The standard deviations were also large, but the quality of answers should not be attributed to the number of words.

TABLE 5. RESPONDENTS' WORD QUANTITIES

Words/Answer	Q1-3 Requirements	Q4-6 Technology	Q7-9 Commonality	Q10-12 Concurrency	Q13-15 Partnering	Respondent Total Avg
JPO (*Manager*)	252 +/- 158	169 +/- 94	124 +/- 57	162 +/- 94	196 +/- 184	903 +/- 451
Mil (*Customer*)	418 +/- 288	340 +/- 281	180 +/- 121	270 +/- 154	313 +/- 375	1520 +/- 853
LM (*Provider*)	200 +/- 106	147 +/- 83	174 +/- 144	157 +/- 74	104 +/- 116	782 +/- 400
Total	**114 +/- 98**	**89 +/- 104**	**55 +/- 49**	**75 +/- 64**	**83 +/- 120**	**1250 +/- 828**

Note. JPO = JSF Program Office, Mil = military services, LM = Lockheed Martin, Q = question, Avg = average

Table 6 shows the breakdown of family codes by the type of respondent—manager (JPO), customer (Services), or provider (LM). Out of 1,564 respondent codes, customers from the Services provided the most codes, but they were proportional to being the largest group of respondents. Customer respondents' codes were about 2.5 times more than the managers from the JPO and four times more than the providers from LM. These proportions held throughout the acquisition strategy challenges' codes, indicating consistency in coding. Although the proportional relationship of total coding versus the family codes of cost, schedule, and performance did not stand up as well, the coding proportions within cost, schedule, and performance were consistent. However, for all coding, remaining objective was challenging for the researcher and required an iterative process of recoding to ensure that the respondents were actually commenting about an acquisition strategy challenge in its relation to one or more of the pillars of acquisition. It is interesting to point out that Balancing Requirements and Harnessing Technology had the most negative codes, with 91 and 89. However, Balancing Requirements and Demanding Commonality had the most positive codes, with 62 and 56. It is interesting that all the acquisition strategy challenges have nearly the same number of recommendations coded—about 50. Within the pillars, the positive categories of Cost Decrease and Cost Decrease Reason were coded the most with 79 and 77 codes. On the negative side, Performance Reduced and Cost Increase Reason had the highest codes with 68 and 52.

TABLE 6. RESPONDENTS' FAMILY CODE BREAKDOWN

	Manager Program Office (10)	*Customer* Services (25)	*Provider* Contractors (7)	Totals (42)
Balancing Requirements Intent (BRI)	13	37	12	62
Balancing Requirements Negative (BRN)	21	54	16	91
Balancing Requirements Positive (BRP)	7	26	11	44
Balancing Requirements Recommendation (BRR)	9	30	8	47
Balancing Requirements Sub-total	49	144	43	236
Technology Intent (TI)	9	23	7	39
Technology Negative (TN)	15	61	13	89
Technology Positive (TP)	3	21	7	31
Technology Recommendation (TR)	15	30	5	50
Harnessing Technology Sub-total	40	134	30	204
Commonality Intent (CMI)	14	28	14	56
Commonality Negative (CMN)	7	39	3	49
Commonality Positive (CMP)	19	32	12	63
Commonality Recommendation (CMR)	9	24	4	37
Demanding Commonality Sub-total	41	118	30	204
Concurrency Intent (CCI)	13	31	6	50
Concurrency Negative (CCN)	16	46	7	69
Concurrency Positive (CCP)	9	28	8	45
Concurrency Recommendation (CCR)	13	31	6	50
Evoking Concurrency Sub-total	49	133	26	189
Partnering Intent (PTI)	12	33	5	50
Partnering Negative (PTN)	10	42	4	56
Partnering Positive (PTP)	14	40	9	63
Partnering Recommendation (PTR)	7	32	4	43
Encouraging Partnering Sub-total	41	143	20	204
Cost Increase (CI)—bad	16	23	6	45
Cost Increase Reason (CIR)—why	22	22	8	52
Cost Decrease (CD)—good	14	44	21	79

TABLE 6. RESPONDENTS' FAMILY CODE BREAKDOWN, CONTINUED

	Manager Program Office (10)	*Customer* Services (25)	*Provider* Contractors (7)	Totals (42)
Cost Decrease Reason (CDR)—why	32	32	13	77
Cost Recommendation (CR)	2	2	0	4
Cost Sub-total	86	120	48	254
Schedule Lengthened (SL)—bad	16	25	10	51
Schedule Lengthened Reason (SLR)—why	13	21	7	41
Schedule Shortened (SS)—good	8	19	7	34
Schedule Shortened Reason (SSR)—why	4	4	0	8
Schedule Recommendation (SR)	3	0	1	4
Schedule Sub-total	43	69	25	137
Performance Reduced (PRD)—bad	14	50	4	68
Performance Reduced Reason (PRR)—why	8	19	3	30
Performance Improved (PI)—good	4	13	3	20
Performance Increased Reason (PIR)—why	3	5	1	9
Performance Recommendation (PR)	2	5	0	7
Performance Sub-total	31	90	11	132
Grand Totals	380	951	233	1564

Note. (XX) = number of respondents

Similar to Figure 2, Figure 4 shows three questionnaire respondents' co-occurrences/c-coefficient tables generated by ATLAS.ti QDA software; these were explained in depth by the dissertation this article is based upon. The table on the left shows the associations between acquisition strategy challenges and the pillars of acquisition. The middle table shows the inter-associations within acquisition strategy challenges. The table on the right shows the interassociations within the pillars of acquisition. For each table, each coded component shows the number of codes in parentheses along the header rows and columns. At each intersection, the co-occurrences and c-coefficients are paired. The intent here is to show some specific examples to eventually discern some obvious general trends. Data that are covered up by overlapping tables are not significant.

FIGURE 4. RESPONDENTS' CO-OCCURRENCE/C-COEFFICIENT TABLES

	CI (45)	CIR (52)	CD (79)	CDR (79)	CR (4)	SL (51)	SLR (41)	SS (34)	SSR (8)	SR (4)	PRD (68)	PRR (30)	PI (20)	PIR (9)	PR (7)
BRI (62)	2/.02	0	7/.05	7/.05	0	1/.01	0	1/.01	0	0	1/.01	2/.02	1/.01	1/.01* (6)	0
BRN (91)	**13/.11**	**11/.08**	3/.02	1/.01	**0**	**13/.10**	8/.06	0	1/.01* (11)	0	**17/.12**	**10/.09**	0	0	0
BRP (44)	0	0	**8/.07**	8/.07	0	0	0	5/.07	2/.04* (5)	0	1/.01	1/.01	2/.03	0	1/.02* (6)
BRR (47)	0	0	3/.02	1/.01	1/.02* (11)	1/.01	1/.01	1/.01	0	**2/.04* (11)**	2/.02	2/.01	0	1/.02* (5)	1/.02* (6)
TI (39)	5/.05	1/.01	1/.01	0	0										
TN (89)	**13/.11**	**12/.09**	0	3/.02	0										
TP (31)	1/.01	1/.01	3/.03	1/.01	0										
TR (50)	2/.02	0	3/.02	0	1/.02* (12)										
CMI (56)	0	0	**11/.09**	**17/.15**	0										
CMN (49)	4/.04	3/.03	4/.03	1/.01	0										
CMP (45)	0	1/.01	**21/.17**	**14/.11**	1/.02* (15)										
CMR (37)	2/.03	0	5/.05	1/.01	0										
CCI (50)	0	0	5/.04	7/.06	0										
CCN (69)	6/.06	**16/.15**	2/.01	0	1/.01* (17)										
CCP (45)	3/.03	2/.02	2/.02	5/.04	0										
CCR (50)	1/.01	0	1/.01	0	**2/.04* (12)**										
PTI (50)	0	0	**10/.08**	**14/.12**	0										
PTN (56)	4/.04	**8/.08**	2/.02	0	0										
PTP (63)	0	0	9/.07	**18/.15**	0										
PTR (43)	0	0	0	1/.01	0										

	BRI (62)	BRN (91)	BRP (44)	BRR (47)	TI (39)	TN (89)	TP (31)	TR (50)	CMI (56)	CMN (49)	CMP (45)	CMR (37)
BRI (62)	0	7/.05	5/.05	0	2/.02	0	1/.01	0	**11/.1**	1/.01	3/.02	4/.04
BRN (91)	-	0	3/.02	5/.04	0	8/.05	0	0	1/.01	**12/.09**	1/.01	2/.02
BRP (44)	-	-										
BRR (47)	-	-										
TI (39)	-	-										
TN (89)	-	-										
TP (31)	-	-										
TR (50)	-	-										
CMI (56)	-	-										
CMN (49)	-	-										
CMP (45)	-	-										
CMR (37)	-	-										
CCI (50)	-	-										
CCN (69)	-	-										
CCP (45)	-	-										

	CI (45)	CIR (52)	CD (79)	CDR (79)	CR (4)	SL (51)	SLR (41)	SS (34)
CI (45)	0	1/.01	2/.02	0	0	**31/.48**	1/.01	1/.01
CIR (52)	-	0	0	0	0	3/.03	**19/.26**	0
CD (79)	-	-	0	0	1/.01* (19)	5/.04	0	**19/.2**
CDR (79)	-	-	-	0	0	0	2/.02	0
CR (4)	-	-	-	-	0	0	0	0
SL (51)	-	-	-	-	-	0	1/.01	0
SLR (41)	-	-	-	-	-	-	0	0
SS (34)	-	-	-	-	-	-	-	0
SSR (8)	-	-	-	-	-	-	-	-
SR (4)	-	-	-	-	-	-	-	-
PRD (68)	-	-	-	-	-	-	-	-

Note. The left table shows strategies versus pillars. The middle table shows strategies' interassociations. The right table shows pillars' interassociations. Data that are covered up by overlapping tables are not significant.

(n), for row/column headers = number of occurrences that were coded, n/c within table where n = co-occurrence and c = c-coefficient—bolded represents strong association if c > .08 and at least two co-occurrences, "*" represents a c-coefficient that may require further investigation—bolded if ratio (r) > 10 and at least two co-occurrences, BRI = Balancing Requirements Intent, BRN = Balancing Requirements Negative, BRP = Balancing Requirements Positive, BRR = Balancing Requirements Recommendation, TI = Technology Intent, TN = Technology Negative, TP = Technology Positive, TR = Technology Recommendation, CMI = Commonality Intent, CMN = Commonality Negative, CMP = Commonality Positive, CMR = Commonality Recommendation, CCI = Concurrency Intent, CCN = Concurrency Negative, CCP = Concurrency Positive, CCR = Concurrency Recommendation, PTI = Partnering Intent, PTN = Partnering Negative, PTP = Partnering Positive, PTR = Partnering Recommendation

The table on the left of Figure 4 has acquisition strategy challenges' component codes depicted in the left column and the pillars of acquisition component codes depicted in the top row. There were several more strong associations circled for this analysis lane as compared to the Reports'/Journals' lane, and each strong association was explained in depth in the dissertation this article is based upon. For the green circle in the middle-left—14 of the 21 co-occurrences between CMP (Commonality Positive) and CD (Cost Decreased) believed there has been and will continue to be cost savings. For the red circle just below, all 16 co-occurrences between CCN (Concurrency Negative) and CIR (Cost Increase Reason) understood that all the early jets, around a 100, need significant and costly modifications that exceeded expectations. For the orange circle to the right between CCR (Concurrency Recommendation) and CR (Cost Recommendation), the c-coefficient required further investigation and yielded one of the best recommendations from a manager, "Maintain program discipline up front, don't over-promise" in the ability to maintain planned concurrency. For the second green circle at the bottom, 14 of the 18 co-occurrences between PTP (Partnering Positive) and CDR (Cost Decrease Reason) reiterated the savings of having multiple countries investing in development and reducing the price per unit. The remaining four occurrences pointed out that Partner countries were given economic opportunities to produce or maintain significant aspects of the F-35.

In the middle table, acquisition strategy challenges interassociations, for the red circle, 8 of the 12 co-occurrences between BRN (Balancing Requirements Negative) and CMN (Commonality Negative) emphasized the resulting differences and uncompromising tradeoffs between the F-35 variants in terms of range requirements, fuel capacities, and weights. In the

right table, pillars of acquisition interassociations, for the green circle, the 19 co-occurrences between CD (Cost Decreased) and SS (Schedule Shortened) read more like revisionist history that has not occurred yet—seven of 19 co-occurrences praised commonality based on future expectations that commonly designed parts will pay off huge dividends in lower sustainment costs for 40 more years of the program and that common mission systems (mission planning and sensors) will reduce the need for Services and other partner nations to look for expensive alternatives. Another seven co-occurrences opined that balancing requirements would have produced earlier savings if no changes in requirements were made. Although three co-occurrences believed that concurrency is currently successful in reducing cost and time, its current reputation is quite the opposite. For the red circle, out of the 31 co-occurrences between CI (Cost Increase) and SL (Schedule Lengthened), 12 were attributed to immature technology, 10 to balancing requirements, three to commonality, four to concurrency, and two to partnering—all five of the acquisition strategy challenges the 2015 study focused on.

> ***...acquisition strategies have both aided and hindered joint aircraft programs' schedule, cost, and performance in terms of Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency, and Encouraging Partnering.***

Answering the Research Questions

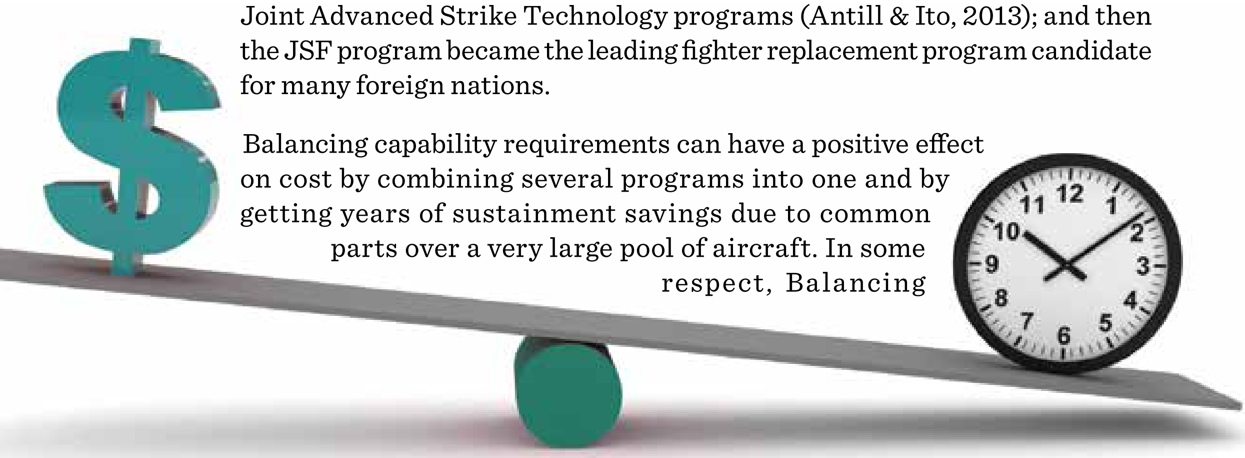
In answering the main research question, the 2015 study helped us understand to a deeper and richer level that acquisition strategies have both aided and hindered joint aircraft programs' schedule, cost, and performance in terms of Balancing Requirements, Harnessing Technology, Demanding Commonality, Evoking Concurrency, and Encouraging Partnering. The following sections will answer each subquestion (SQ) about the acquisition strategy's original intent, its positive and negative effects on schedule, cost, and performance, and what improvements could be made. Individual acquisition strategy challenge taxonomies will show relationships between strategies and pillars, and emphasize relativeness of what was learned from each analysis lane from the 2015 study, and together will help summarize lessons learned.

For each acquisition strategy challenge's summary taxonomy, there will be contributing factors and arrows in the appropriate color matching to the previous synthesized Figure 1 or to the previous co-occurrence/c-coefficient tables in Figures 2 and 4. The previous joint aircraft programs' analysis lane will be depicted in brown, government and think tank reports and scholarly journals will be in purple, and respondents' questionnaires will be in black. The components for acquisition strategy challenges and pillars of acquisition will usually be in red/negative or green/positive boxes if they emanated from a strong association starting with an acquisition strategy challenge. Otherwise, some component boxes will remain clear. Within the tan circle of acquisition strategy challenges, for strong relationships to other strategies, only contributing factors from the analysis lane from which they came from will be shown in the appropriate colored text. Within the olive area, contributing factors between pillars of acquisition components will be shown in white text over the appropriate colored arrow to also indicate from which analysis lane it came.

SQ1 on Balancing Requirements

The original intent of balancing requirements, especially with joint aircraft programs, is to merge and trade off the needs of several U.S. military services and nations into one weapon system, thus combining several programs into one in order to save money and time as opposed to each Service or country developing its own aircraft program. The result of Balancing Requirements would be a weapon system that has commonality in design, structure, mission systems, and parts among possible variants. Specifically, the JSF program initially combined eight U.S. military services' programs: Advanced Short Take-Off/Vertical Landing, the Short Take-Off/Vertical Landing Strike Fighter, the Common Affordable Lightweight Fighter, the Multi-Role Fighter, the Advanced Tactical Aircraft, the Naval Advanced Tactical Fighter, the Advanced Attack/Advanced/Fighter Attack, and the Joint Advanced Strike Technology programs (Antill & Ito, 2013); and then the JSF program became the leading fighter replacement program candidate for many foreign nations.

Balancing capability requirements can have a positive effect on cost by combining several programs into one and by getting years of sustainment savings due to common parts over a very large pool of aircraft. In some respect, Balancing



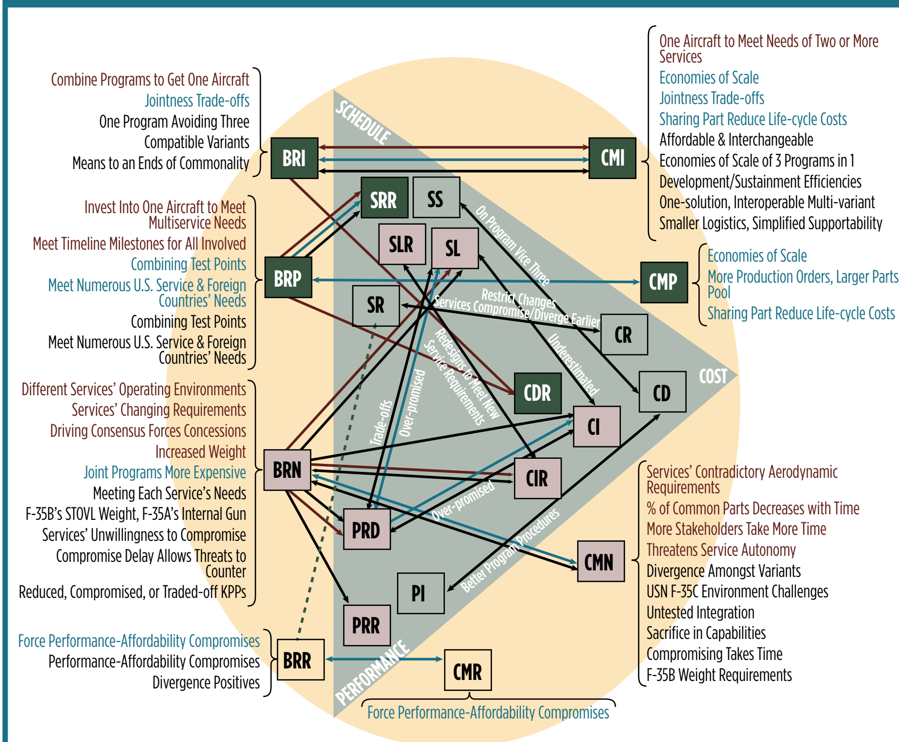
Requirements reduces schedule because of the one single design effort and the consolidation of test points between variants. On the other hand, balancing capability requirements has had a negative effect on the JSF schedule because it has taken longer than expected for the Services to agree on requirements since they operate in different environments. Although commonality is great in the long term for parts replacement, in the short term, the JSF program took more time and money to design those common parts within specified weight criteria, especially to meet the requirements of the USMC's F-35B, which took a SWAT (STOVL Weight Attack Team) to solve. During the time it took for the Services to compromise and the engineers to design, Services updated and added key performance parameters (KPP) as the threat evolved, thereby increasing cost.

To improve the process of balancing capability requirements, one can either force performance-affordability compromises between the Services earlier and more often or be willing to diverge from the designs between the variants earlier. As we learned from questionnaire respondents only, initial airframe commonality expectation was 80%, and this was reduced in actuality to 30 to 40% by 2015. However, respondents highlighted that mission systems and avionics between all variants were between 95 and 100%. The USMC gave up time, but still got their STOVL replacement to the AV-8 Harrier. The USMC did not care as much about low observability as the USAF did, so much so that the USMC ditched the internal gun for an external gun pod that increases the radar cross section. Although the USAF's desire for an internal gun conflicted directly with the USMC's need for a vertical lift fan in almost the same location of the aircraft, the USAF did lose speed and range performance to incorporate other aspects that were needed for USMC STOVL engine capability.

As a matter of relativeness between the analysis lanes, Figure 5 shows a summary taxonomy of all the Balancing Requirements components' contributing factors and how they affect the pillars of acquisition and how the pillars of acquisition react to one another under the guise of balancing requirements. We learned from previous joint aircraft programs (in brown), especially from the 1960s TFX F-111, that combining multiple Services' requirements is a difficult challenge. The difference with the JSF program is that the USMC needs the F-35B more now than the USN needed the F-111 back then. Numerous negative contributing factors, mostly from the previous joint aircraft program analysis (in brown) and from the respondent questionnaire analysis (in black), are more prevalent than positive contributing factors, especially in relation to cost, which RAND concluded cost more for joint programs than for multiple single-Service programs

(Lorell et al., 2013). The relationship that stands out within the pillars of acquisition is that overpromised performance, coupled by underestimated cost and schedule, resulted in Services' capability tradeoffs.

FIGURE 5. BALANCING REQUIREMENTS SUMMARY TAXONOMY



Note. CI = Cost Increased, CIR = Cost Increase Reason, CD = Cost Decreases, CDR = Cost Decrease Reason, CR = Cost Recommendation, SL = Schedule Lengthened, SS = Schedule Shortened, SSR = Schedule Shortened Reason, SR = Schedule Recommendation, PRD = Performance Reduced, PRR = Performance Reduced Reason, PI = Performance Improved, BRI = Balancing Requirements Intent, BRN = Balancing Requirements Negative, BRP = Balancing Requirements Positive, BRR = Balancing Requirements Recommendation, CMI = Commonality Intent, CMN = Commonality Negative, CMP = Commonality Positive, CMR = Commonality Recommendation

Red = Previous Joint Aircraft Programs, Blue = Government/Think Tank Reports and Scholarly Journals, Black = Respondents' Questionnaires

SQ2 on Harnessing Technology

The original intent of harnessing technology is for the weapon system to give warfighters an asymmetric advantage over their potential adversaries. Although potential adversaries have developed low-observable (LO) cruise missiles, the United States is considered the leader in aircraft LO

development. That development started with the F-117 stealth fighter and continues with the B-2 stealth bomber and the F-22 fighter—which are being slowly countered by potential adversaries. As reported in annual DOT&E reports (2004-2014), the F-35 fighter is expected to be the most advanced and dominant aircraft in the world for decades to come, not just for its next generation LO stealth technology, but for its fused and integrated sensors, weapons, and electronic attack used in conjunction with LO.

Harnessing technology is not cheap, and it takes time, but it eventually improves tactical performance. Even with COTS technology, the military services eventually manipulate the technology to meet military specifications much like what happened to the 1990s JPATS T-6. In conjunction with concurrency, numerous advanced technologies, even if somewhat immature, can be brought to the warfighter quicker by flying with the systems sooner to fix problems earlier so aircrews can have the chance to survive against an advanced threat. What makes the F-35 and the F-22 "5th Gen" is not just its low observability, but its fusion of sensors. The F-35 uses LO stealth structure, active-electronically scanned array (AESA) radar, and electronic warfare that has incrementally improved over the last 20 years. Although air forces have used helmets with displays in the visor for over 15 years, the JSF's helmet mounted display system (HMDS) is so advanced that you cannot employ the F-35 without it, and if it fails, the pilot returns to base. The latest and greatest technologies are the electro-optical targeting system (EOTS; forward looking air-to-air and air-to-ground infrared search and track system) and the distributed aperture system (a 360-degree day/night electro-optical system), which give the pilot a protective sphere of situational awareness of incoming aircraft and missile threats.

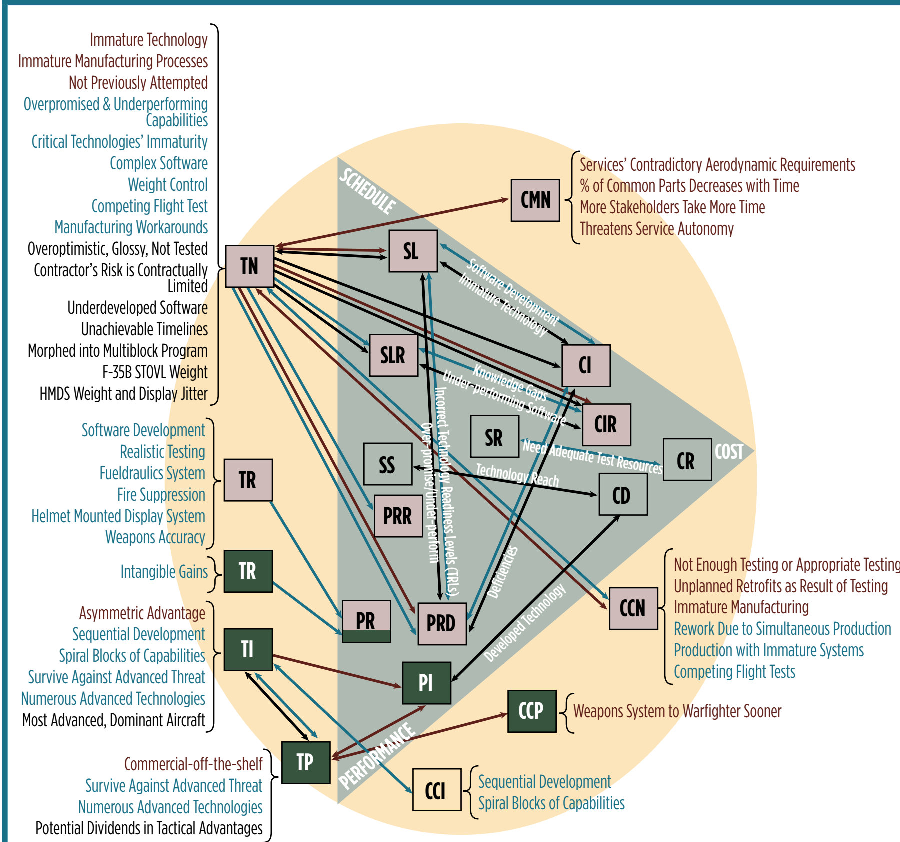
Making sure these technologies work as advertised usually takes more time and money than planned. Much like the 1960s TFX F-111 variable wing sweep (for high and low speeds) and variable inlet (for the new engine) and the 1980s JVX V-22 tilt rotor technology challenges, the JSF F-35 experienced many technology challenges just to fit the unique lift fan engine design for the USMC F-35B STOVL into the same basic airframe mid-section for all F-35 variants (DOT&E, 2010)—part of the negativeness of commonality. Not only was the F-35B overweight by several hundred pounds for quite some time before a solution was found (Blickstein et al., 2011), but the HMDS, which is exactly the same for all variants, has also been too heavy by ounces to be safe during ejection, and had a jittery display that was fixed separately (GAO, 2014). Concerning academics courseware for pilot and maintenance training and for aircraft maintenance tracking records, ALIS is still immature and requires workarounds that use legacy

academic courseware and legacy maintenance-tracking systems (Everstine, 2015). The root causes to all these late schedules and cost increases are due to the complex and immature software that is still being developed, written, and tested. Many of the requirements' developers within each Service were overly optimistic about the envisioned capabilities and the time and cost it would take to deliver. As much as concurrency helped this endeavor, more redesigns and rework than expected occurred during initial flight training.

Most of the improvements recommended for the JSF were shown in Table 2 to correct technical deficiencies. Unclosed items include software development, the HMDS, weapons accuracy due to AESA and EOTS issues, fueldraulics' survivability, and fire suppression. All of these were pursued for the USAF's IOC declaration in August 2016, five years later than planned (Insinna, 2016; Pike, 2012). On the positive side, for any great leaps in technology for the United States to maintain tactical advantages in aerial warfare, we just have to try—because even temporary failures produce gains in knowledge.

As a matter of relativeness between the analysis lanes, Figure 6 shows a summary taxonomy of all the Harnessing Technology components' contributing factors and how they affect the pillars of acquisition and how the pillars of acquisition react to one another under the guise of "mature" technology. Figure 6 is visually significant in showing that all analysis lanes (in brown, purple, and black) contributed several factors to Technology Negative (TN) technology that had numerous negative effects on schedule, cost, and performance. As for recommendations, the green/positive and red/negative Technology Recommendations (TR) point to a split Performance Recommendation (PR). Within the pillars of acquisition olive triangle, although there may be a chance that new technologies can significantly improve performance and then reduce cost and decrease schedule, it is more likely that these new technologies will be overpromised, will underperform, and will result in deficient capabilities due to immaturity with underdeveloped software and knowledge gaps. Government program managers need to do a better job in assessing technology readiness levels of new systems.

FIGURE 6. HARNESSING TECHNOLOGY SUMMARY TAXONOMY



Note. CI = Cost Increased, CIR = Cost Increase Reason, CD = Cost Decreased, CR = Cost Recommendation, SL = Schedule Lengthened, SLR = Schedule Lengthened Reason, SS = Schedule Shortened, SR = Schedule Recommendation, PRD = Performance Reduced, PRR = Performance Reduced Reason, PI = Performance Improved, TI = Technology Intent, TN = Technology Negative, TP = Technology Positive, TR = Technology Recommendation, CMN = Commonality Negative, CCI = Concurrency Intent, CCN = Concurrency Negative, CCP = Concurrency Positive

Red = Previous Joint Aircraft Programs, Blue = Government/Think Tank Reports and Scholarly Journals, Black = Respondents' Questionnaires

SQ3 on Demanding Commonality

The original intent of demanding commonality is to drive down development, manufacturing, and sustainment costs by having a high level of interchangeable parts and software for a weapon system to reap the benefits of economies of scale when large quantities are bought. Commonality was the result of Balancing Requirements among three Services into a single, multivariant aircraft that has about 40% airframe commonality, but

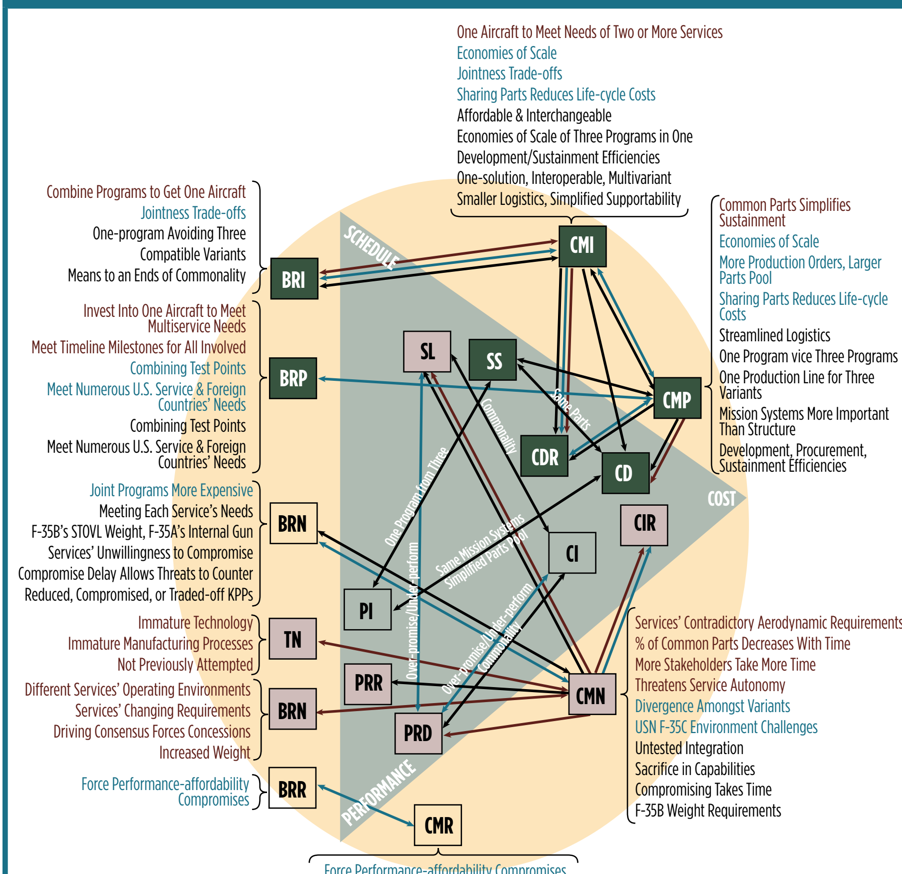
almost 100% commonality in mission systems, sustainment, and training software, eventually supporting over 3,000 F-35s. Although the airframe commonality among the F-35A/B/C is not as high as the 80% expected, there is still just one manufacturing line that can produce all three variants, significantly lowering the price per unit, according to annual GAO reports, which became attractive to foreign nations. Furthermore, with only 40% airframe commonality, most of the parts that can be replaced are common and that will save hundreds of millions of dollars over the decades to come. As more F-35s are built, over time there are less spare parts per aircraft needed, which drives logistical costs down. Even with partner-unique parts, the commonality of the computerized maintenance and training system based on the ALIS backbone will also save hundreds of millions of dollars once ALIS has its full capability. Furthermore, beyond the airframe, with almost 100% common mission and training systems, combining tests points shortens the schedule.

Similar to the Balancing Requirements previous discussion in terms of negativeness, in order to meet requirements within the different Services' operating environments, compromise entailed making sacrifices in KPPs. For example, the USAF accepted a less than optimal ejection system with a backwards canopy and less fuel capacity to meet USMC F-35B weight needs. When compromise was not attained, divergence occurred. For example, the USAF F-35A maintained the internal gun, while the USMC F-35B removed the internal gun to make more room for the lift fan part of their engine, and the USN F-35C has a third larger wing than the F-35As and Bs. These separate designs contradict the ideal of demanding commonality. A common recommendation from the managers and the providers is to keep forcing performance-affordability compromises early in the program, but even the customers admitted that Services' autonomy and parochialism usually win at a cost of development time.

As a matter of relativeness between the analysis lanes, Figure 7 shows a summary taxonomy of all the Demanding Commonality components' contributing factors, and how they affect the pillars of acquisition, and how the pillars of acquisition react to one another under the pressure of commonality. Figure 7 is a visual "draw" between the positives and negatives, looking very similar to those of Balancing Requirements in Figure 5. Most of the positive factors and relationships, however, came from the questionnaire responses (in black), and the negatives came from previous joint aircraft programs (in purple). If commonality was graded over time, it has been losing early and often, evident by a couple of Nunn-McCurdy breaches, but GAO reports and several respondents believe there is high probability that

over time the story will change to be a successful strategy. Within the pillars of acquisition olive triangle, Demanding Commonality was sold as a cost- and time-savings strategy to make three Service fighter programs into one, while the same mission systems and simplified parts pool will eventually reap a cost benefit. However, the expectations of these savings were oversold in the near term, and the Services will not let you forget that they sacrificed on some KPPs to attempt to meet the commonality goals.

FIGURE 7. DEMANDING COMMONALITY SUMMARY TAXONOMY



Note. CI = Cost Increased, CIR = Cost Increase Reason, CD = Cost Decreased, CDR = Cost Decrease Reason, SL = Schedule Lengthened, SS = Schedule Shortened, PRD = Performance Reduced, PRR = Performance Reduced Reason, PI = Performance Improved, BRI = Balancing Requirements Intent, BRN = Balancing Requirements Negative, BRP = Balancing Requirements Positive, BRR = Balancing Requirements Recommendation, CMI = Commonality Intent, CMN = Commonality Negative, CMP = Commonality Positive, CMR = Commonality Recommendation

Red = Previous Joint Aircraft Programs, Blue = Government/Think Tank Reports and Scholarly Journals, Black = Respondents' Questionnaires

SQ4 on Evoking Concurrency

The original intent of evoking concurrency was to systematically overlap development, production, test, and fielding of a weapon system to get it quickly to the warfighters with some capability in order to improve follow-on production lots based on what was learned from the warfighters, thus saving time and money. Concurrency was offered as a strategy to answer the 1980s Packard Commission about how to prevent long and costly defense acquisition programs (Eide & Allen, 2012). In theory, concurrency shortens the program by having test and training conducted simultaneously. It should also force the program to account for supportability and logistics earlier. Several respondents explained how concurrency was planned for 4 to 6 years over three blocks of aircraft—Blocks 1, 2, and 3, but the program ended up with 11 different blocks—Blocks 0, 1A, 1B, 2A, 2B, 3i, 3F, 4.1, 4.2, 4.3, and 4.4. By 2011, the JPO's plan for U.S. F-35s was to retrofit 25% of the fleet based on procuring 600 aircraft before the end of initial operational test and evaluation in March 2012 (Blickstein et al., 2011). In comparison, the F-111 produced 141 F-111s for the USAF before changes could be made on the manufacturing line for a 25% concurrency rate against 547 total USAF F-111s (Coulam, 1977; Richey, 2005). The final concurrency rate will not be known until after the fact.

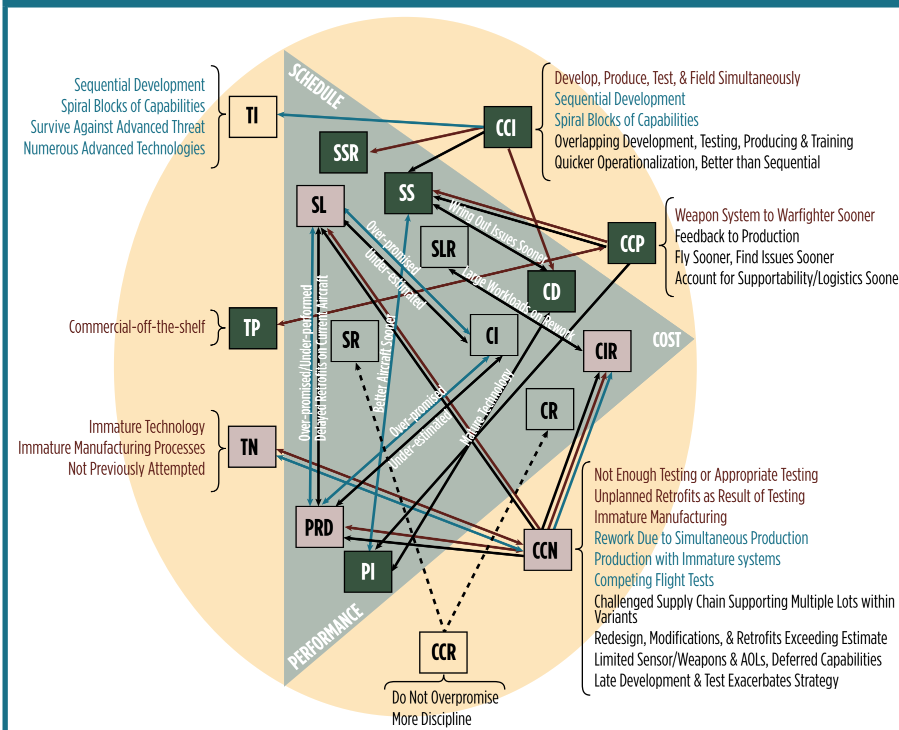
The 1980s JVX V-22 relied heavily on concurrency. It entered full-rate production without ensuring mature manufacturing processes that required a redesign and retrofit of hydraulics and electrics for the tilt-rotor system, resulting in a Nunn-McCurdy breach in 2001, after the USMC had begun MV-22 aircrew training in 2000 (General Accounting Office, 2003; Whittle, 2010). Seven years later, the USMC declared MV-22 IOC in 2007 (Whittle, 2010). There is little argument that with concurrency more F-35 issues have been found earlier than otherwise predicted, and many issues have been corrected quicker with cut-ins on the production line, but the amount and complexity of the retrofits exceeded expectations. Most of this disappointment is due to incomplete, late flight testing and overpromised capabilities. Although the USAF started training in 2012 with an aircraft of very limited capability, right after the second and last Nunn-McCurdy breach, IOC will not be declared until 2016. Instead of dealing with three blocks of aircraft during these 4 years, one "operator" respondent emphasized that pilots and maintainers will have dealt with seven blocks—Block 0.5, 1A, 1B, 2A, 2B, 3i (small "i" for initial combat capability), and 3F. Follow-on F-35 development will include four more blocks starting in 2019 with one every 2 years—Block 4.1, 4.2, 4.3, and 4.4 (Tirpak, 2015).

> ***One provider said this about JSF concurrency: "The original model for JSF is solid. Build test aircraft, build training aircraft while testing is going on, and then accelerate to full rate production when the aircraft has matured sufficiently."***

Unfortunately, Concurrency Recommendations (CCR) did not have any strong associations for the Reports'/Journals' co-occurrence/c-coefficient tables in Figure 2, but the left table for the Respondents' co-occurrence/c-coefficient tables in Figure 4 showed they required further investigation. It suggested that program managers should be more disciplined when accepting immature technology and be more suspicious of the promises of capabilities early on. By digging deeper into the respondents' questionnaires, some appropriate recommendations were found. A few manager respondents suggested ensuring the technology readiness levels are more accurate, reserving concurrency for a less complex weapon system, and to never accept this level of concurrency again. Several customer respondents suggested making the early production lots smaller in quantity. One provider said this about JSF concurrency: "The original model for JSF is solid. Build test aircraft, build training aircraft while testing is going on, and then accelerate to full rate production when the aircraft has matured sufficiently."

As a matter of relativeness between the analysis lanes, Figure 8 shows a summary taxonomy of all the Evoking Concurrency components' contributing factors, and how they affect the pillars of acquisition, and how the pillars of acquisition react to one another under the pressure of commonality. Factors contributing to Concurrency Negative (CCN) are more apparent than to Concurrency Positive (CCP). All analysis lanes (previous in brown, reports/journals in purple, and questionnaires in black) significantly contributed to the concurrency negative impacts on schedule, cost, and performance. Within the pillars of acquisition olive triangle, there are dueling positive and negative component triangles between schedule, cost, and performance. The positive Schedule Shortened (SS)-Cost Decreased (CD)-Performance Improved (PI) triangle assumes mature technology and expects to wring out problems sooner to get a better aircraft sooner. The negative Schedule Lengthened (SL)-Cost Increased (CI)-Performance Reduced (PRD) triangle harps on overpromised results and underperforming reality that was reiterated by the reports/journal (in purple) and the respondents' questionnaire analyses (in black).

FIGURE 8. EVOKING CONCURRENCY SUMMARY TAXONOMY



Note. AOL = Aircraft Operational Limits, CI = Cost Increased, CIR = Cost Increase Reason, CD = Cost Decreased, CR = Cost Recommendation, SL = Schedule Lengthened, SS = Schedule Shortened, SSR = Schedule Shortened Reason, SR = Schedule Recommendation, PRD = Performance Reduced, PI = Performance Improved, TI = Technology Intent, TN = Technology Negative, TP = Technology Positive, CCI = Concurrency Intent, CCN = Concurrency Negative, CCP = Concurrency Positive, CCR = Concurrency Recommendation

Red = Previous Joint Aircraft Programs, Blue = Government/Think Tank Reports and Scholarly Journals, Black = Respondents' Questionnaires

SQ5 on Encouraging International Partnering

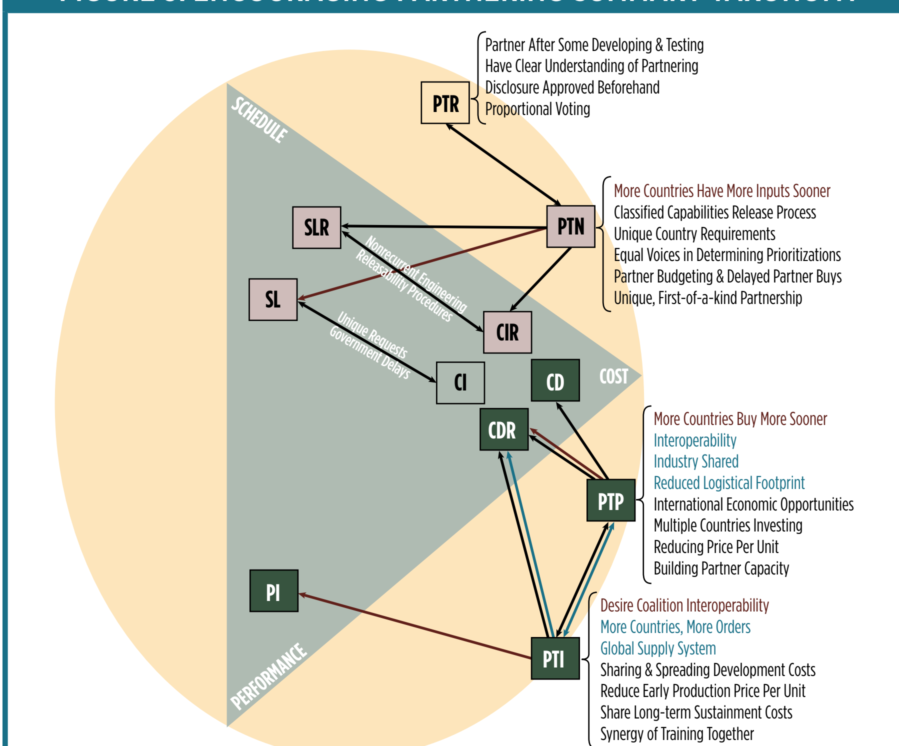
The original intent of encouraging partnering was to get partner nations invested early in the program to help pay for development and to order and receive aircraft early in the program. This would reduce the price per unit and result in coalition interoperability. In return for taking risk by investing in development, the Partners get to learn how to develop a 5th Gen aircraft that no other opportunity would provide. Over the life of the 40-year program, sustainment costs are also shared. As a business model to defend against U.S. congressional budget cutters, international support goes a long way, especially when the expectation is to train and fight together.

Most respondents referred to building partner capacity as the positive by-product of partnering—it promotes the United States' interest in not fighting alone, in fighting on the same level on the same side, and in spreading the burden of policing the world. Besides reducing U.S. development and production costs and building partnership capacity in the expectation we will fight together, partner nations reap the added benefits of economic opportunity in their own countries as their companies become suppliers for LM. As one of the first partners, U.K.'s Martin-Baker will manufacture all the F-35A/B/Cs' ejection system, representing billions of dollars in revenue to outfit over 3,000 potential aircraft (JPO, 2014). Italy gained 10,000 jobs created by their industrial participation in the program including the activities of the Cameri Final Assembly Check-Out, Maintenance, Repair, Overhaul, and Upgrade (MRO&U) plant to assemble their 90 F-35s via license from LM (Marrone, 2013). Even Turkey was approved to be the first of three nations to oversee MRO&U for the F135 engine (Butler, 2015). Partner countries also get to bid on managing spare parts logistics hubs for Europe and Asia (Butler, 2015).

This arrangement, guided by the JPO (2007) *Production, Sustainment, and Follow-on Development* (PSFD) *Memorandum of Understanding* (MOU) signed by the partner nations listed in the MOU's Table 3 (Chapter 2), is a first-of-a-kind partnership that comes with several challenges not experienced in normal FMS. In FMS cases, the weapon system is already developed, there is extra capacity to produce and to train if requested, and the United States usually dictates the terms. In this partnership, no matter how many aircraft a nation is purchasing, the foreign countries get an equal vote right from the start, and when priorities cannot be agreed upon, it takes longer. If a country has a unique requirement, although it pays for it alone, it takes nonrecurring engineering time away from other design and test priorities. But the most difficult challenge concerning the strong relationship between cost and schedule is the problem of getting classified capabilities released to partner nations. This is especially important when USAF and foreign F-35A student pilots will be training with one another, taught by one another by USAF and Partner instructors, and flying each other's aircraft at Luke AFB—a process known as "pooling." Although there are DoD offices in charge of protecting critical technologies, five customer respondents commented that other disclosure, security, and exportability offices have not fully embraced or understood the PSFD MOU's intent. Improvements to remedy these negatives come strictly from respondents' questionnaires and start with recommending delaying partnering until development is complete, settling disclosure issues before partners invest, and making voting proportional to the number of aircraft ordered.

As a matter of relativeness between the analysis lanes, Figure 9 shows a summary taxonomy of all the Encouraging Partnering components' contributing factors and how they affect the pillars of acquisition and how the pillars of acquisition react to one another under the pressure of commonality. Although Figure 9 looks visually even, the contributing factors to Partnering Positive (PTP) of building partner capacity, interoperability, and true coalition warfare, and the tangible, international economic opportunities brought forward mostly by the reports/journals (in purple) and the respondents' questionnaires (in black), outweigh the procedural frustrations in Partnering Negative (PTN) of releasability and voting that will be solved over time by necessity. However, there will be a price to pay in time and cost within the pillars of acquisition olive triangle. This is especially true with unique country requests and the fact that partner nations all have civilian governments that may delay ordering and purchasing from the previously agreed plan, which instantly raises the price per unit that particular year.

FIGURE 9. ENCOURAGING PARTNERING SUMMARY TAXONOMY



Note. CI = Cost Increased, CIR = Cost Increase Reason, CD = Cost Decreased, CDR = Cost Decrease Reason, SL = Schedule Lengthened, SLR = Schedule Lengthened Reason,

PI = Performance Improved, PTI = Partnering Intent, PTP = Partnering Positive, PTN = Partnering Negative, PTR = Partnering Recommendation

Red = Previous Joint Aircraft Programs, Blue = Government/Think Tank Reports and Scholarly Journals, Black = Respondents' Questionnaires

Final Recommendations

All program managers in the defense industry and in commercial business should note the relations within the pillars of acquisition that were confirmed. Although the JSF program is in the 21st century, lessons on the pillars of acquisition were first relearned in the 20th century with the TFX, JVS, and JPATS programs and reconfirmed by analysis on the JSF program. If technical performance cannot be sacrificed, but is hard to achieve, expect cost to rise and schedule to lengthen. If cost is fixed due to budgeting and the schedule is expected to stay the same, performance will suffer and requirements will be lowered. If a program manager wants to maintain schedule, cost usually goes up for more personnel, but performance may suffer again due to rushing.

What is learned is not necessarily about scandal and tragedy; however, all program managers in the defense industry will recognize familiar lessons and learn some new ones:

- Balancing requirements between the Services is not recommended for large joint MDAPs, especially for aircraft. However, balancing requirements within a single-Service program to meet cost, schedule, and performance goals should be expected. This takes disciplined leadership and the ability to manage expectations with transparency of what can be achieved.

- Harnessing technology, or reaching for immature technology, should be done only if needed to maintain a tactical advantage in a strategic environment. Otherwise, program managers need to assess technology readiness levels better and not be so enamored with contractors' glossy brochures of cost, scheduling, and performance.

- Demanding commonality may be costly and time consuming as the result of balancing requirements in the near term, but should pay off with a streamlined logistics system in the long term. Do not underestimate the commonality of mission

systems (communications, sensors, mission planning, and prognostic health management) and of training systems (courseware and simulators).

- Evoking concurrency as a primary strategy will be disappointing if technology was incorrectly represented as mature, but this is a matter for better discipline, management, and execution, and it cannot be addressed with Flyvbjerg's (2003) megaprojects' paradox that the program is too big to fail. If concurrency occurs because a development program falls behind schedule, project managers need to adjust expectations, but that still requires discipline.
- Encouraging partnering needs to be understood better from the beginning and embraced by government agencies and military services in terms of disclosure, security, and exportability—especially if the United States truly wants to build partner capacity and to have true, interoperable, coalition warfare.

References

Air Education and Training Command. (2010, March). *A training system for the 21st century: JPATS and the T-6A* (AFD-061213-005). Retrieved from https://www.rairarubiabooks.com/related-pdf-enter-trainer-extended-wings.html

Antill, P., & Ito, P. (2013). The UK and the Joint Strike Fighter: The trials and tribulations of international collaborative procurement. *International Journal: Canada's Journal of Global Policy Analysis, 68*(1), 13–29. doi:10.1177/002070201306800103

Art, R. (1969). *The TFX decision: McNamara and the military.* Boston, MA: Little, Brown and Company.

Associated Press. (2016, April 26). *Senator says fighter program has been scandal and tragedy.* Retrieved from http://www.foxnews.com/us/2016/04/26/senator-says-fighter-program-has-been-scandal-and-tragedy.html

Blickstein, I., Boito, M., Drezner, J., Dryden, J., Horn, K., Kallimani, J., … Wong, C. (2011). *Root cause analyses of Nunn-McCurdy breaches: Zumwalt Class Destroyer, Joint Strike Fighter, Longbow Apache, and Wideband Global Satellite (Vol. 1)*. Retrieved from RAND Corporation website: http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1171.1.pdf

Blickstein, I., Nemfakos, C., & Sollinger J. M. (2013). Digging out the root cause: Nunn-McCurdy breaches in major defense acquisition programs. *Defense Acquisition Research Journal, 20*(2), 128–153. Retrieved from http://www.dau.mil/publications/DefenseARJ/ARJ/ARJ66/ARJ_66-Blickstein.pdf

Butler, A. (2015, January 6). F-35 nearing in-service debuts: With a handful of fighters nearing operations, F-35 is still the one to beat. *Aviation Week and Space Technology, 176*(47), 69. Retrieved from http://aviationweek.com/defense/j-10b-f-35-nearing-service-debuts

Coulam, R. (1977). *Illusions of choice: The F-111 and the problem of weapons acquisition reform.* Princeton, NJ: Princeton University Press.

Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Defense Acquisition University. (2006). *Risk management guide for DoD acquisition* (6th ed.). Retrieved from http://www.dau.mil/publications/publicationsDocs/RMG%206Ed%20Aug06.pdf

Department of Defense. (2012). *Joint Primary Aircraft Training System (JPATS) Selected Acquisition Report (SAR)* (RCS: DD-A&T(Q&A)823-560). Retrieved from http://www.globalsecurity.org/military/library/budget/fy2012/sar/jpats_december_2012_sar.pdf

Department of Defense Inspector General. (2013, September 30). *Quality assurance assessment of the F-35 Lightning II program.* Retrieved from http://www.dodig.mil/pubs/documents/DODIG-2013-140.pdf

Director, Operational Test and Evaluation. (2004). *Fiscal year 2004 annual report F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2004/pdf/af/2004F-35JSF.pdf

Director, Operational Test and Evaluation. (2005). *Fiscal year 2005 annual report F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2005/pdf/af/2005f-35.pdf

Director, Operational Test and Evaluation. (2006). *Fiscal year 2006 annual report F-35 Joint Strike Fighter (JSF).* Retrieved from http://www.dote.osd.mil/pub/reports/FY2006/pdf/af/2006f35.pdf

Director, Operational Test and Evaluation. (2007a). *Fiscal year 2007 annual report F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2007/pdf/af/2007f35.pdf

Director, Operational Test and Evaluation. (2007b). *Joint Primary Aircraft Training System (JPATS)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2007/pdf/navy/2007jpats.pdf

Director, Operational Test and Evaluation. (2008). *F-35 Lightning II Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2008/pdf/dod/2008f35.pdf

Director, Operational Test and Evaluation. (2009). *F-35 Lightning II Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2009/pdf/dod/2009f35.pdf

Director, Operational Test and Evaluation. (2010). *F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2010/pdf/dod/2010f35jsf.pdf

Director, Operational Test and Evaluation. (2011). *F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2011/pdf/dod/2011f35jsf.pdf

Director, Operational Test and Evaluation. (2012). *F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2012/pdf/dod/2012f35jsf.pdf

Director, Operational Test and Evaluation. (2013). *F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2013/pdf/dod/2013f35jsf.pdf

Director, Operational Test and Evaluation. (2014). *F-35 Joint Strike Fighter (JSF)*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2014/pdf/dod/2014f35jsf.pdf

Dunne, T. (2011). A case for the F-35 Lightning. *National Defence and the Canadian Armed Forces, 11*(4), 55–60. Retrieved from http://www.journal.forces.gc.ca/vo11/no4/55-dunne-eng.asp

Eide, P. K., & Allen, C. D. (2012). The more things change, acquisition reform remains the same. *Defense Acquisition Research Journal, 19*(1), 99–120. Retrieved from http://www.dau.mil/pubscats/PubsCats/AR%20Journal/arj61/Eide61.pdf

Ergas, H. (2009). Trillions for military technology: How the Pentagon innovates and why it costs so much. *Agenda: A Journal of Policy Analysis and Reform, 16*(2), 95–98. Retrieved from http://search.proquest.com/docview/209796602?accountid=28180

Everstine, B. (2015, April 25). *Problems plaguing F-35's next-gen maintenance system*. Retrieved from http://www.airforcetimes.com/story/military/2015/04/15/problems-facing-f35-maintainers-automated-system/25781075/

Flyvbjerg, B. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge, UK: Cambridge University Press.

Friese, S. (2013). *ATLAS.ti 7 user guide and reference.* Berlin, Germany: ATLAS.ti Scientific Software Development. Retrieved from http://atlasti.com/wp-content/uploads/2014/05/atlasti_v7_manual_201312.pdf?q=/uploads/media/atlasti_v7_manual_201312.pdf

Gantt, M. E. (2002). *Strapping in and bailing out: Navy and Air Force joint acquisition of aircraft*. Retrieved from Defense Technical Information Center website: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA420511

Gates, R. (2009, December 2). Statement on Afghanistan to the Senate Armed Services Committee. Retrieved from http://archive.defense.gov/speeches/default.aspx?mo=12&yr=2009

General Accounting Office. (1973). *F-111 Aircraft*. Retrieved from http://www.gao.gov/assets/80/78812.pdf

General Accounting Office. (1986). *DoD acquisition: Case study of the Navy V-22 Osprey joint vertical lift aircraft program* (GAO-NSIAD-86-45S-7). Retrieved from http://www.gao.gov/assets/210/208624.pdf

General Accounting Office. (1990). *Navy aviation: The V-22—Progress and problems* (GAO/NSIAD-91-45). Retrieved from http://www.gao.gov/assets/220/213295.pdf

General Accounting Office. (1994). *V-22 development—Schedule extended, performance reduced, and costs increased* (GAO/NSIAD-94-14). Retrieved from http://www.gao.gov/assets/220/218986.pdf

General Accounting Office. (1997). *Navy aviation: V-22 cost and capability to meet requirements are yet to be determined* (GAO/NSIAD-98-13). Retrieved from http://www.gao.gov/assets/230/224788.pdf

General Accounting Office. (2003). *Defense acquisitions: Assessments of major weapon programs* (GAO-03-476). Retrieved from http://www.gao.gov/assets/240/238188.pdf

General Accounting Office. (2004). *Defense acquisitions: Assessments of major weapon programs* (GAO-04-248). Retrieved from http://www.gao.gov/assets/250/241798.pdf

Gertler, J. (2011). *V-22 Osprey tilt-rotor aircraft: Background and issues for Congress* (RL31384). Retrieved from http://www.fas.org/sgp/crs/weapons/RL31384.pdf

Government Accountability Office. (2005). *Defense acquisitions: Assessments of selected major weapon programs* (GAO-05-301). Retrieved from http://www.gao.gov/assets/250/245884.pdf

Government Accountability Office. (2006). *Defense acquisitions: Assessments of selected major weapon programs* (GAO-06-391). Retrieved from http://www.gao.gov/assets/250/249468.pdf

Government Accountability Office. (2007). *Defense acquisitions: Assessments of selected weapon programs* (GAO-07-406SP). Retrieved from http://www.gao.gov/assets/260/258416.pdf

Government Accountability Office. (2008). *Defense acquisitions: Assessments of selected weapon programs* (GAO-08-467SP). Retrieved from http://www.gao.gov/assets/280/274156.pdf

Government Accountability Office. (2009). *Defense acquisitions: Assessments needed to address V-22 aircraft operational and cost concerns to define future investments* (GAO-09-482). Retrieved from http://www.gao.gov/assets/290/289395.pdf

Government Accountability Office. (2010). *Defense acquisitions: Assessments of selected weapon programs* (GAO-10-388SP). Retrieved from http://www.gao.gov/assets/310/302379.pdf

Government Accountability Office. (2011a). *Defense acquisitions: Assessments of selected weapon programs* (GAO-11-233SP). Retrieved from http://www.gao.gov/assets/320/317081.pdf

Government Accountability Office. (2011b). *Joint Strike Fighter: Restructuring places program on firmer footing, but progress is still lagging* (GAO-11-677T). Retrieved from http://www.gao.gov/assets/130/126245.pdf

Government Accountability Office. (2012a). *Defense acquisitions: Assessments of selected weapon programs* (GAO-12-400SP). Retrieved from http://www.gao.gov/assets/590/589695.pdf

Government Accountability Office. (2012b). *Joint Strike Fighter: Restructuring added resources and reduced risk, but concurrency is still a major concern* (GAO-12-525T). Retrieved from http://www.gao.gov/assets/590/589454.pdf

Government Accountability Office. (2012c). *Joint Strike Fighter: Restructuring added resources and reduced risk, but concurrency is still a major concern* (GAO-12-525T). Retrieved from http://www.gao.gov/assets/590/589454.pdf

Government Accountability Office. (2013). *Defense acquisitions: Assessments of selected weapon programs* (GAO-13-294SP). Retrieved from http://www.gao.gov/assets/660/653379.pdf

Government Accountability Office. (2014). *Defense acquisitions: Assessments of selected weapon programs* (GAO-14-340SP). Retrieved from http://www.gao.gov/assets/670/662184.pdf

Government Accountability Office. (2015). *Defense acquisitions: Assessments of selected weapon programs* (GAO-15-342SP). Retrieved from http://www.gao.gov/assets/670/668986.pdf

Hunter, A. (1998). *TSR2 with hindsight*. Brighton, U.K.: RAF Historical Society.

Insinna, V. (2016, August 2). Air Force declares F-35A ready for combat. *Defense News.* Retrieved from http://www.defensenews.com/story/breaking-news/2016/08/02/f35-ioc-air-force-operational-acc-combat/87948142/

Jogulu, U., & Pansiri, J. (2011). Mixed methods: A research design for management doctoral dissertations. *Management Research Review 34*(6), 687–701. doi:10.1108/01409171111136211

Joint Program Office. (2007). *JSF production, sustainment, and follow-on development memorandum of understanding*. Retrieved from http://www.jsf.mil/downloads/documents/JSF_PSFD_MOU_-_07_Feb_07.pdf

Logan, D. (1998). *General Dynamics F-111 Aardvark*. Atglen, PA: Schiffer.

Leedy, P., & Ormrod, J. (2009). *Practical research: Planning and design* (9th ed.). Upper Saddle River, NJ: Pearson.

Lorell, M. A., Kennedy, L., Leonard, R. S., Munson, K., Abramzon, S., An, D. L., & Guffey, R. A. (2013). *Do joint fighter programs save money?* Retrieved from Rand Corporation website: http://www.rand.org/content/dam/rand/pubs/monographs/MG1200/MG1225/RAND_MG1225.pdf

Marrone, A. (2013). Italy and the F-35: Rationales and costs. *International Journal: Canada's Journal of Global Policy Analysis, 68*(1), 31-48. doi:10.1177/002070201306800104

McCullough, A. (2015, July 20). First ospreys for Japan. *Air Force Magazine: Online Journal of the Air Force Association*. Retrieved from http://www.airforcemag.com/DRArchive/Pages/2015/July%202015/July%2015%202015/First-Ospreys-for-Japan.aspx

Melese, F., Franck, R., Angelis, D., & Dillard, J. (2007). Applying insights from transaction cost economics to improve cost estimates for public sector purchases: The case of U.S. military acquisition. *International Public Management Journal, 10*(4), 357–385. doi:10.1080/10967490701683511

Miller, J. (1982). *General Dynamics F-111 "Aardvark."* Fallbrook, CA: Aero Publishers.

Moustakas, C. (1990). *Heuristic research: Design, methodology, and applications.* Newbury Park, CA: Sage.

Patton, M. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Pike, J. (2011). *F-35 Joint Strike Fighter (JSF) commonality*. Retrieved from the GlobalSecurity.Org website: http://www.globalsecurity.org/military/systems/aircraft/f-35-commonality.htm

Pike, J. (2012). *F-35 Joint Strike Fighter (JSF) Lightning II program*. Retrieved from the GlobalSecurity. Org website: http://www.globalsecurity.org/military/systems/aircraft/f-35-program.htm

President's Blue Ribbon Commission on Defense Management. (1986). *A quest for excellence—final report to the President*. Retrieved from http://usacac.army.mil/cac2/CSI/docs/Gorman/06_Retired/01_Retired_1985_90/07_86_PackardCommission_FinalReport/01_PackardCommission_FinalReport.pdf

Richey, G. (2005). *F-111 systems engineering case study*. Wright Patterson AFB, OH: Air Force Institute of Technology. Retrieved from https://acc.dau.mil/CommunityBrowser.aspx?id=37599&lang=en-US

Shank, G. (2006). *Qualitative research: A personal skills approach*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Summers, R. (1965). *Cost estimates as predictors of actual weapons costs: A study of major hardware articles* (RM-3061-PR). Retrieved from RAND Corporation website: http://www.rand.org/pubs/research_memoranda/RM3061.html

TFX contract investigation, parts 1-10: Hearings before the Permanent Subcommittee on Investigations of the Committee on Government Operations, Senate, 88th Cong. 1 (1963). Washington, DC: U.S. Government Printing Office.

TFX contract investigation, parts 1-3: Hearings before the Permanent Subcommittee on Investigations of the Committee on Government Operations, Senate, 91st Cong. 2 (1970). Washington, DC: U.S. Government Printing Office.

Tirpak, J. A. (2015, June 1) F-35 block 4 taking shape. *Air Force Magazine: Online Journal of the Air Force Association*. Retrieved from http://www.airforcemag.com/DRArchive/Pages/2015/June%202015/June%2001%202015/F-35-Block-4-Taking-Shape.aspx

Uda, R. T. (2012). *Exploring factors contributing to improvised explosive device (IED) terror campaigns in the United States homeland* (Publication No. 3530099). Available from Dissertations & Theses @ Northcentral University; ProQuest Dissertations & Theses Global (1085492122). Retrieved from http://search.proquest.com.proxy1.ncu.edu/docview/1085492122?accountid=28180

Wacker, J. (1998). A definition of theory: Research guidelines for different theory-building research methods in operations management. *Journal of Operations Management, 16*(4), 361–385. doi:10.1016/S0272-6963(98)00019-9

Wicht, A. C., & Crawley, E. F. (2012). Relieving joint pain: Planning government acquisition of complex common systems. *Defense Acquisition Research Journal, 19*(3), 221–248. Retrieved from http://www.dau.mil/pubscats/PubsCats/AR%20Journal/arj64/Wicht_Crawley_ARJ63.pdf

Whittle, R. (2010). *The dream machine: The untold history of the notorious V-22 Osprey*. New York, NY: Simon & Schuster.

Yin R. (1994). *Case study research: Designs and methods* (2nd ed.). Newbury Park, CA: Sage.

Young, H., Vorwig, P., Lieberman, M., Wigginton, W., House, M., Black, W., & Maravilla, C. (2010). Aerospace and defense industries. *The International Lawyer, 44*(1), 337–353. Retrieved from http://search.proquest.com/docview/645692834?accountid=28180

Zikmund, W., Babin, B., Carr, J., & Griffin, M. (2010). *Business research methods.* (8th ed.). Mason, OH: Thomas/South-Western.

Biography

Col Roger Witek, USAF (Ret.), is currently the Director of Training and Services for Tapestry Solutions. Dr. Witek was chief of Air Education and Training Command's Joint Strike Fighter Coordination and Training Office, responsible for the pilot and maintenance training programs for the USAF and seven Partner nations. He has a doctorate in business administration, is a former vice-wing and squadron commander, and is a graduate of the USAF Weapons Instructor Course and the School of Advanced Air and Space Studies. As a Secretary of Defense Corporate Fellow, he produced a research report, "Hard Skills, Soft Skills, Savviness, and Discipline—Recommendations for Successful Acquisition" (available at http://dcmo.defense.gov/Portals/47/Documents/Products_Services/SecDef_Fellows/08_09__Boeing_Report.doc).

(E-mail address: rcwitek1989@gmail.com)


Image designed by Diane Fleischer

Federally Mandated FURLOUGHS:

The Effect on **Organizational Commitment** and **Organizational Citizenship Behavior**

Robert L. Shepherd

As the U.S. Congress searches for ways to remain proficient, while cutting expenditures in an effort to decrease the federal deficit, federal government employee furloughs could become more widely implemented. In order to sustain productivity within the U.S. Government federal civilian service while implementing furloughs, supervisors could benefit from a better understanding of how furloughs affect the organizational commitment (OC) and organizational citizenship behavior (OCB) of federal government employees. The purpose of this study was to determine the effect of federally mandated furloughs on the OC and OCB of federal government employees. To answer this question, a nonexperimental survey study was conducted on federal government civilian employees and active duty military employees. By identifying the negative effects associated with furloughs, managers can more effectively address such things as decreases in morale, lowered productivity rates, and increases in employee turnover.

DOI: *http://dx.doi.org/10.22594/dau.15-740.24.01*
Keywords: *retention, motivation, sequestration, layoffs, productivity*

Introduction to the Problem

Furloughs in private companies and government organizations have become more frequent since the U.S. economic downturn of 2008 (Brutocao & Marshall, 2011; Halbesleben, Wheeler, & Paustian-Underdahl, 2013; Shannon, 2010). The organizational effects of furloughs, including both the positive and negative impacts associated with them, have been studied (Halbesleben et al., 2013; Hohman, Packard, Finnegan, & Jones, 2013; Lee & Sanders, 2013; Osborne, Smith, & Huo, 2013; Shannon, 2010). Federal civilian service furloughs implemented by the U.S. Government often are short-term, such as those that occur when Congress misses a budget deadline or fails to pass a continuing resolution. Because of the unique culture of federal agencies and the potential for federally mandated involuntary furloughs to occur in the future, additional research on how furloughed employees' organizational commitment (OC) and organizational citizenship behavior (OCB) change in response to implemented furloughs could prove beneficial to supervisors.

Researchers question the long-term effects of furloughs and argue that the implementation of furloughs can be counterproductive (Duggan, Lewis, & Milluzzi, 2010; Halbesleben et al., 2013). Two examples of this counterproductivity are reduced employee efficiency and diminished employee morale (Badiru, 2014). OC is perpetuated by the employees' level of desire to stay with that organization (Elias, 2007). If furloughs have a negative effect on OC, the outcome may lead to higher turnover rates and more expenditure for hiring and training new employees.

OCB is an employee's actions contributing to and supporting a larger social and psychological atmosphere within the organization (Bolino, Klotz, Turnley, & Harvey, 2013). Employees with high OCB go above the general requirements of their job, often going out of their way to help others within the organization. If furloughs have a negative effect on the OCB of employees, the organization

could suffer lower morale and productivity because employees have lowered their contribution to a larger social and psychological atmosphere within the organization.

Kelman (2006) maintained that differences in human attitudes justify the need to identify differences among human circumstances across organizations. Government employees tend to have a greater sense of job security and feel more resentment in response to a breach of their psychological contract. Organizations, to remain competitive, must continuously alter organizational culture both to adapt to and to accept change (Cristian-Liviu, 2013). As more research is accomplished using federal government employees as participants, more results will become available on how federal employees react in different situations based not only on circumstances and attitudes but on varying kinds of internal organizational cultures.

Definition of Terms

Federal Service Organization

For this study, the term "federal service organization" will refer to an agency within the federal government that currently employs federal service employees. The federal government organization under study contained both active duty military personnel and federally employed civilians.

Organizational Commitment

For the purpose of this study, the term "organizational commitment" (OC) will refer to the employee's belief in the values and mission of the organization and willingness to exert extra effort to support the success of the organization. Mowday, Steers, and Porter (1979) defined three aspects of OC. The first aspect is a strong connection with, and acceptance of, organizational values and goals. The second aspect is that employees who exhibit strong OC want to exert an extended amount of effort for the organization.

The third aspect involves an employee's strong desire to remain an active member of their organization. The three aspects of OC manifest in employee behaviors in devotion to their work and their desire to see the organization succeed. For the purpose of this study, the behaviors and attitudes of the employees under study, in relationship to the organization for which they work, will determine the depth of the employees' OC. The OC portion of this

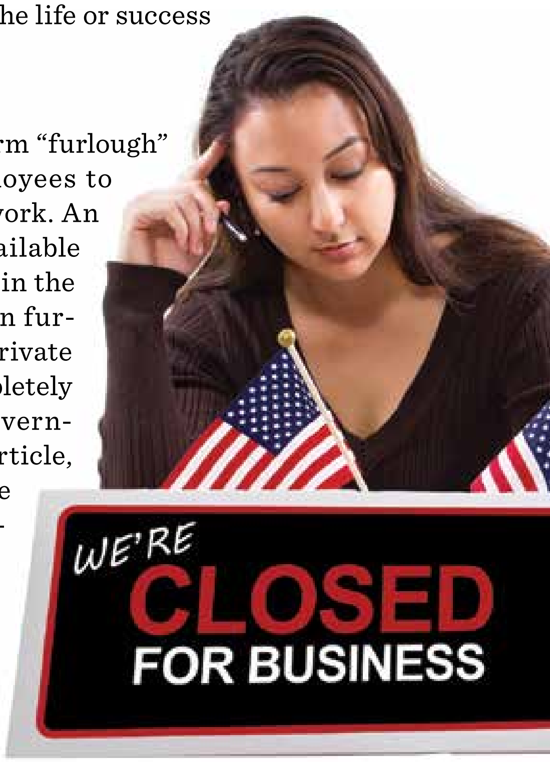
study will investigate the employees' willingness to accept any job assignment to stay with the organization and the employees' emotional connection to the fate of the organization.

Organizational Citizenship Behavior

For the purpose of this study, the term "organizational citizenship behavior" (OCB) is defined as an employee's willingness to exert effort above their formal job description to support the goals and values of the organization. Gurbuz (2009) gave an in-depth definition of OCB as behavior that the organization does not require but that happens as a result of an individual's free will to help others progress or achieve a task. He identified five components of OCB: altruism, conscientiousness, sportsmanship, courtesy, and civic virtue. Altruism is a selfless act done to help others. Conscientiousness is the act of effectively using time and going beyond the normal requirements of the role as defined by the normal behaviors within the organization. Sportsmanship is the act of enduring or accepting the difficulties of the work without complaint. Courtesy is the willingness to share information with others to prevent problems related to work. Civic virtue is the employee's contribution to the life or success of the organization.

Furlough

For the purpose of this study, the term "furlough" will refer to the act of requiring employees to take an unpaid leave of absence from work. An examination of the limited research available on the effects of furloughs reveals a gap in the current body of knowledge. Research on furloughs falls exclusively in the areas of private companies and state governments, completely excluding employees of the federal government. In the instance studied for this article, U.S. Government federal civilian service employees were subjected to both scheduled furloughs due to sequestration and unscheduled furloughs due to an unexpected government shutdown.

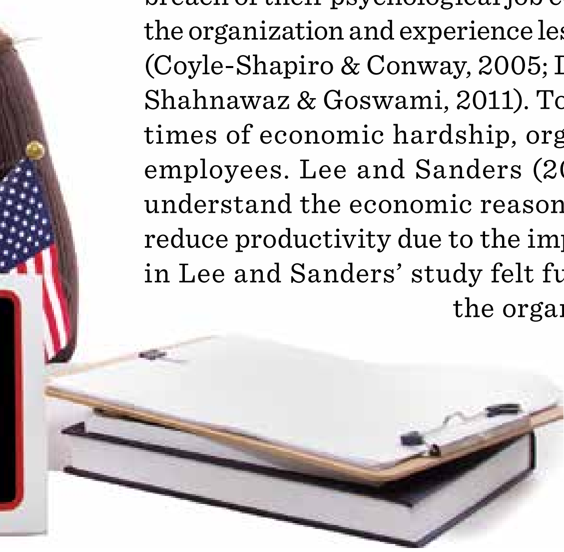


Background of the Study

Private and public companies made more use of furloughs to reduce expenditures in 2010 than they have since World War II (Shannon, 2010). During March 2013, the President of the United States implemented sequestration to reduce the deficit by $80.5 billion (Ferrell, 2014). In response to sequestration, the Secretary of Defense ordered mandatory administrative furloughs for 624,404 civilians employed by the Department of Defense (DoD). Both military personnel and contractors associated with the DoD were exempt from furloughs. DoD civilian employees were initially required to take 11 unpaid days off work. The number of furlough days was reduced to 6 unpaid days during the months between July and September 2013.

Current research conducted on OC focuses on leader-member exchange theory and job satisfaction (Ariani, 2012). A positive correlation between leader-member exchange theory and OC suggests that strong leader-member relationships foster higher OC among employees. Employees who exhibit strong OC are more likely to stay with the organization during periods of distress such as organizational change (Elias, 2007).

Although OCB is not a formal job requirement, the behavior often has a positive effect on the functionality of the organization (Bolino et al., 2013). Recent research on OCB reveals that OC, job satisfaction, and trust are leading determinates of OCB (Gurbuz, 2009). If employees feel there is a breach of their psychological job contract, they could be less committed to the organization and experience less job satisfaction and lower productivity (Coyle-Shapiro & Conway, 2005; DelCampo, 2007; Nadin & Cassell, 2007; Shahnawaz & Goswami, 2011). To sustain trust among employees during times of economic hardship, organizations must be honest with their employees. Lee and Sanders (2013) maintained that employees who understand the economic reasoning behind furloughs are less likely to reduce productivity due to the implementation of furloughs. Participants in Lee and Sanders' study felt furloughs were the best option for both the organization and themselves because their leadership communicated to them the fiscal state of their organization.



Purpose of the Study

The purpose of this survey study was to test the theory that employee OCB and OC are affected by furloughs among federal government employees. Participants were drawn from two groups: active duty military personnel who did not experience an involuntary furlough and federal civilian service employees who did experience an involuntary furlough. The first group of participants was exempt from furloughs. Participants in the second group were subjected to planned furloughs as a result of mandated federal government sequestration and unplanned furloughs as a result of a government shutdown in 2013 between October 1 and October 16. Furloughed government employees did receive back pay for the unplanned furlough period even though they did not work during the time of the government shutdown. Furloughed employees did not receive any pay for the six scheduled furlough days in 2013. Surveys were administered to both samples and were statistically compared. By recognizing how employees change their behavior in reference to furloughs, supervisors might be able to decrease the negative psychological effects associated with increased stress and the emotional exhaustion employees incur as a result of furloughs (Hohman et al., 2013).

Research Questions

Research question 1

How do furloughs implemented in a federal government agency affect employee OC?

H_O: Furloughs in a federal government agency have no statistically significant effect on employee OC.

H_O: $\mu_1 = \mu_2$ Where μ_1 is the mean of nonfurloughed employees' OC and μ_2 is the mean of furloughed employees' OC.

H_A: Furloughs in a federal government agency have a statistically significant effect on employee OC.

H_A: $\mu_1 \neq \mu_2$ Where μ_1 is the mean of nonfurloughed employees' OC and μ_2 is the mean of furloughed employees' OC.

Research question 2

How do furloughs implemented in a federal government agency affect employee OCB?

H_O: Furloughs in a federal government agency have no statistically significant effect on employee OCB.

H_O: $\mu_1 = \mu_2$ Where μ_1 is the mean of nonfurloughed employees' OCB and μ_2 equals the mean of furloughed employees' OCB.

H_A: Furloughs in a federal government agency have a statistically significant effect on employee OCB.

H_A: $\mu_1 \neq \mu_2$ Where μ_1 is the mean of nonfurloughed employees' OCB and μ_2 equals the mean of furloughed employees' OCB.

Methodology

The methodological approach for this study was a nonexperimental survey-based quantitative design. The quantitative research methodology allowed the researcher to remain objective, not biasing the results of the

data with his perspective. Online surveys allowed the researcher to gather the data using participants from a federal government agency without the use of government time or equipment.

Sample Size

The sample size was calculated using the G*power3 Scientific Power Calculator (Faul, Erdfelder, Lang, & Buchner, 2007). The researcher calculated the sample size for an independent sample two-tailed t-test using the effect size, alpha error, power level, and allocation ratio. The effect size used to calculate the sample size was .8. The sample size was calculated using an alpha error of .05 and the power level used was .95. The allocation ratio was 1, and the population size for the agency under study was 209. The total sample size calculated for the study was 84, with 42 furloughed participants and 42 nonfurloughed participants.

The organization granted permission to conduct the study on December 13, 2013. Organizational permission was granted with the contingency that employees were not allowed to complete the survey at work during work time or using government equipment. The study yielded a total of 87 combined responses from active duty military and federal civilian service employees. The nonfurloughed group represented by active duty military members yielded a total of 43 responses with 42 usable responses. The furloughed group represented by federal civilian service employees yielded 44 total responses with 43 usable responses. The total number of usable completed responses collected was 85.

Instrumentation/Measures

Organizational Commitment Questionnaire. The instrument used to measure the OC of participants for this study was the OC questionnaire developed by Mowday et al. (1979). The OC questionnaire consists of nine

statements that participants rated on a seven-point Likert scale reflecting how much they agreed with the statement. The OC questionnaire contained nine positively worded items (Mowday et al.). The scale consisted of 1 = "Strongly Disagree," 2 = "Moderately Disagree," 3 = "Slightly Disagree," 4 = "Neither Agree or Disagree," 5 = "Slightly Agree," 6 = "Moderately Agree," and 7 = "Strongly Agree." A score of 63 on the instrument indicates the highest OC and a score of 9 indicates the lowest OC. The OC questionnaire is an open-source survey instrument and does not require permission for use.

Organizational Citizenship Behavior Scale. The instrument used to measure OCB in this study was the OCB scale developed by Smith, Organ, and Near (1983). The OCB scale consists of 16 items with three questions negatively worded and reverse scored. The instrument uses a five-point Likert scale: 1 = "Never," 2 = "Seldom," 3 = "Occasionally," 4 = "Often," and 5 = "Almost Always." The three negatively worded and reverse-scored items were "take undeserved work breaks," "coast toward the end of the day," and "a great deal of time spent with personal telephone conversations."

The OCB scale measures two dimensions of OCB: altruism and general compliance. Seven items on the scale are used to determine altruism, and the remaining nine items are used to measure general compliance. A score of 80 on the OCB scale indicates maximum OCB, and a score of 16 indicates minimum OCB. The researcher obtained permission to use the OCB scale from Dr. Janet Near (personal communication, August 14, 2013).

Data Analysis

Organizational Commitment. Data were analyzed using a two-tailed independent sample *t*-test. The nonfurloughed group contained 42 usable surveys completed by active duty participants. The furloughed group consisted of 43 usable surveys completed by civil service employees. The first test used the OC questionnaire to calculate a significance value between the two samples in Statistical Package for the Social Sciences (SPSS) version 21. Rejection of the null hypothesis occurred if the significance value of the two means was less than or equal to the alpha value of .05. Failure to reject the null hypothesis occurred if the significance value of the two means was greater than the probability value of .05.

Organizational Citizenship Behavior. To determine the statistical difference of OCB between the furloughed and nonfurloughed federal government employees, the data were analyzed using a two-tailed independent samples *t*-test for combined OCB, altruism, and general compliance. The OCB scale was used to calculate a significance value between the two samples in SPSS version 21. Three areas of OCB were analyzed using the OCB scale.

The first area examined by the OCB scale was the combined scores, which indicated the overall OCB for furloughed and nonfurloughed employees. The second area was the altruism of furloughed and nonfurloughed federal government employees. The third area was general compliance for furloughed and nonfurloughed employees. Rejection of the null hypothesis occurred if the significance value of the two means was less than or equal to the probability value of .05. Failure to reject the null hypothesis occurred if the significance value of the two means was greater than the probability value of .05.

Results

Organizational Commitment

SPSS version 21 was used to calculate the means, standard deviation, and standard error mean of data collected from furloughed and nonfurloughed federal government employees using the OC questionnaire (Table 1). The mean score for nonfurloughed employees was 5.198 and for furloughed employees was 4.494. The standard deviation for nonfurloughed employees was 1.244 and for furloughed employees was 1.358. The standard error of the mean for nonfurloughed employees was .192 and for furloughed employees was .207.

TABLE 1. ORGANIZATIONAL COMMITMENT GROUP STATISTICS

Group	*N*	Mean	Standard Deviation	Standard Error Mean
Nonfurlough	42	5.198	1.244	.192
Furlough	43	4.494	1.358	.207

An independent samples *t*-test was calculated at a confidence interval of 95% in SPSS version 21 using the data collected with the OC questionnaire. The results of the two-tailed independent samples *t*-test are presented in Table 2. The *p*-value calculated for OC between furloughed and nonfurloughed employees, with equal variance assumed, was .015, indicating $p < .05$. “Sig (2-tailed)” represents the *p*-value in Table 2. The effect size was calculated using Cohen’s *d*. The effect size calculated for the OC portion of this study was .548, indicating a medium effect size.

TABLE 2. ORGANIZATIONAL COMMITMENT INDEPENDENT SAMPLES *t*-TEST

Levene's Test for Equality of Variance	Sig	*t*	*df*	Sig (2-tailed)	Mean Difference	Standard Error Difference	95% Confidence Interval of the Difference		Cohen's *d*
							Lower	Upper	
Equal Variance Assumed	.157	2.494	83	.015*	.7049	.2826	.1428	1.2669	.548
Equal Variance Not Assumed		2.497	82.665	.015	.7049	.2823	.1433	1.266	

Note. $^{*}p < .05$

The results of the study indicate the difference in OC between furloughed and nonfurloughed employees is statistically significant and does not support the null hypothesis. The results of the independent samples *t*-test performed for OC indicate furloughs have a statistically significant effect on the OC of federal civilian service employees.

Organizational Citizenship Behavior

SPSS version 21 was used to calculate the means, standard deviation, and standard error mean of data collected from furloughed and nonfurloughed federal government employees using the OCB scale (Table 3). The mean score for nonfurloughed employees was 3.787 and for furloughed employees was 3.493. The standard deviation for nonfurloughed employees was .563 and for furloughed employees was .448. The standard error of the mean for nonfurloughed employees was .087 and for furloughed employees was .068.

TABLE 3. ORGANIZATIONAL CITIZENSHIP BEHAVIOR GROUP STATISTICS

Group	*N*	Mean	Standard Deviation	Standard Error Mean
Nonfurlough	42	3.787	.563	.087
Furlough	43	3.493	.448	.068

An independent samples *t*-test was calculated at a confidence interval of 95% in SPSS version 21 using the data collected with the OCB scale. The results of the two-tailed independent samples *t*-test are presented in Table 4. The *p*-value calculated for OCB between furloughed and nonfurloughed employees, with equal variance not inferred, was .009, indicating $p < .05$. "Sig (2-tailed)" represents the *p*-value in Table 5. The effect size was calculated using Cohen's *d*. The effect size calculated for the OCB portion of this study was .579, indicating a medium effect size.

TABLE 4. ORGANIZATIONAL CITIZENSHIP BEHAVIOR INDEPENDENT SAMPLES *t*-TEST

Levene's Test for Equality of Variance	Sig	*t*	*df*	Sig (2-tailed)	Mean Difference	Standard Error Difference	95% Confidence Interval of the Difference		Cohen's *d*
							Lower	Upper	
Equal Variance Assumed	.038	2.672	83	.009	.2945	.1102	.0753	.5137	.579
Equal Variance Not Assumed		2.665	78.190	.009*	.2945	.1104	.0744	.5114	

Note. $^*p < .05$

The results of the study indicate the difference in OCB between furloughed and nonfurloughed employees is statistically significant and fails to support the null hypothesis. The results of the independent samples *t*-test performed for OCB indicate furloughs have a statistically significant effect on the OCB of federal government employees.

Organizational Citizenship Behavior: Altruism

The researcher then used the same method to analyze only the items on the OCB scale designed to measure altruism. SPSS version 21 was used to calculate the means, standard deviation, and standard error mean of collected data pertaining to altruism from furloughed and nonfurloughed federal government employees using the OCB scale (Table 5). The mean score for nonfurloughed employees was 3.582 and for furloughed employees was 3.259. The standard deviation for nonfurloughed employees was .737 and for furloughed employees was .609. The standard error mean for nonfurloughed employees was .114 and for furloughed employees was .093.

TABLE 5. ORGANIZATIONAL CITIZENSHIP BEHAVIOR ALTRUISM GROUP STATISTICS

Group	*N*	Mean	Standard Deviation	Standard Error Mean
Nonfurlough	42	3.582	.737	.114
Furlough	43	3.259	.609	.093

An independent samples *t*-test was calculated at a confidence interval of 95% in SPSS version 21 using the data collected with the OCB scale altruism items. The results of the two-tailed independent samples *t*-test are presented in Table 6. The *p*-value calculated for OCB altruism items between furloughed and nonfurloughed employees, with equal variance not inferred, was .031, indicating $p > .05$. “Sig (2-tailed)” represents the *p*-value in Table 6. The effect size was calculated using Cohen’s *d*. The effect size calculated for the OCB altruism portion of this study was .477, indicating a medium effect size.

TABLE 6. ORGANIZATIONAL CITIZENSHIP BEHAVIOR ALTRUISM INDEPENDENT SAMPLES *t*-TEST							
t	***df***	**Sig (2-tailed)**	**Mean Difference**	**Standard Error Difference**	**95% Confidence Interval of the Difference**		**Cohen's *d***
					Lower	**Upper**	
2.197	79.425	.031*	.3225	.1468	.0304	.6146	.477

Note. *$p > .05$

The results of the study indicated the difference in altruism between furloughed and nonfurloughed federal government employees is not statistically significant and supports the null hypothesis. The results of the independent samples *t*-test performed for OCB altruism indicate furloughs do not have a statistically significant effect on the altruism characteristic of OCB of federal civilian service employees.

Organizational Citizenship Behavior: General Compliance

The researcher analyzed only the items on the OCB scale designed to measure general compliance using a two-tailed independent sample *t*-test. SPSS version 21 was used to calculate the means, standard deviation, and standard error mean of collected data pertaining to general compliance from furloughed and nonfurloughed federal government employees using the OCB scale (Table 7). The mean score for nonfurloughed employees was 3.747 and for furloughed employees was 3.674. The standard deviation for nonfurloughed employees was .517 and for furloughed employees was .458. The standard error mean for nonfurloughed employees was .080 and for furloughed employees was .070.

TABLE 7. ORGANIZATIONAL CITIZENSHIP BEHAVIOR GENERAL COMPLIANCE GROUP STATISTICS				
Group	***N***	**Mean**	**Standard Deviation**	**Standard Error Mean**
Nonfurlough	42	3.747	.517	.080
Furlough	43	3.674	.458	.070

An independent samples *t*-test was calculated at a confidence interval of 95% in SPSS version 21 using the data collected with the OCB scale general compliance items. The results of the two-tailed independent samples *t*-test are presented in Table 8. The *p*-value calculated for OCB general compliance items between furloughed and nonfurloughed employees, with equal variance not inferred, was .012, indicating $p < .05$. "Sig (2-tailed)" represents the *p*-value in Table 8. The effect size was calculated using Cohen's *d*. The effect size calculated for the OCB general compliance portion of this study was .558, indicating a medium effect size.

TABLE 8. ORGANIZATIONAL CITIZENSHIP BEHAVIOR GENERAL COMPLIANCE INDEPENDENT SAMPLES *t*-TEST

t	*df*	Sig (2-tailed)	Mean Difference	Standard Error Difference	95% Confidence Interval of the Difference		Cohen's *d*
					Lower	Upper	
2.572	81.285	.012*	.2727	.1060	.0617	.4836	.558

Note. *$p < .05$

The results of the study indicated the difference in general compliance between furloughed and nonfurloughed federal government employees is statistically significant and does not support the null hypothesis. The results of the independent samples *t*-test performed for OCB general compliance indicate furloughs have a statistically significant effect on the general compliance characteristic of OCB of federal government employees.

Discussion

Organizational Commitment

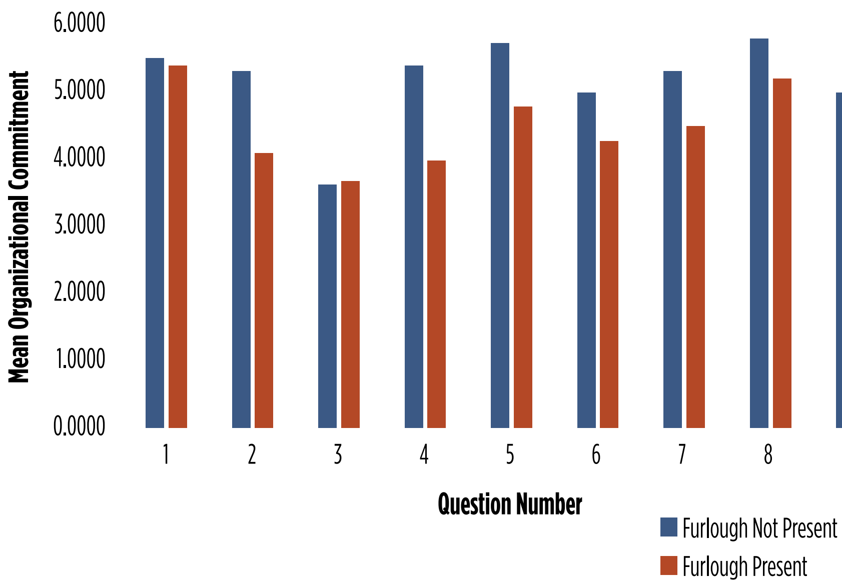
The results of the OC questionnaire revealed an average mean score of 4.494 for furloughed federal government employees and 5.198 for nonfurloughed federal government employees. The OC of furloughed employees was significantly lower than nonfurloughed employees. A closer look at the data reveals that nonfurloughed employees ranked higher in OC on all items listed on the survey except item number three (Figure 1).

Item number three on the OC questionnaire states, "I would accept almost any type of job assignment in order to keep working for this organization." Active duty military members representing the nonfurloughed sample of the study may have indicated that they would not accept almost any other type of job assignment in order to continue working for their current organization because they did not have as much time invested in their careers as did civilian employees making up the furloughed sample for the study. Most civilian employees participating in the study were retired from active duty military service. Many of the civilian employees who participated in the study had obtained civilian jobs in career fields similar to those they had when they worked in the military. In addition, active duty military members tend to be more mobile, moving to a new duty station every 4 years on the average. This mobility gives the active duty military employees more flexibility in terms of changing organizations within the same branch of the military when compared to civilian workers.

Employees who were required to take furlough days may have felt less security within their jobs, fearing more permanent measures to cut expenditures such as downsizing.

The highest occurrence of differences between furloughed and nonfurloughed federal government employees in terms of OC were items number two and four on the questionnaire. Item number two states, "I talk up the organization to my friends as a great organization to work for." Item number four states, "I find that my values and the organization's values are very similar." The fact that both of these items scored significantly lower among furloughed employees when compared to nonfurloughed employees could indicate animosity towards the organization as a result of furloughs. Employees who were required to take furlough days may have felt less security within their jobs, fearing more permanent measures to cut expenditures such as downsizing. Furloughed employees may have felt there was no justification for furloughs and that their values were no longer in line with organizational values.

FIGURE 1. MEAN SCORES FOR ORGANIZATIONAL COMMITMENT



1. I am willing to put in a great deal of effort beyond that normally expected in order to help this organization be successful.
2. I talk up this organization to my friends as a great organization to work for.
3. I would accept almost any type of job assignment in order to keep working for this organization.
4. I find that my values and the organization's values are very similar.
5. I am proud to tell others that I am part of this organization.
6. This organization really inspires the very best in me in the way of job performance.
7. I am extremely glad that I chose this organization to work for over others I was considering at the time I joined.
8. I really care about the fate of this organization.
9. For me this is the best of all possible organizations for which to work.

Organizational Citizenship Behavior

The results of the OCB scale indicated an average mean score of 3.493 for furloughed federal government employees and 3.787 for nonfurloughed federal government employees. The average mean score for furloughed employees was significantly lower than for nonfurloughed employees in the federal government agency under investigation (Table 3). Of the 16 items on the OCB scale, furloughed employees ranked higher than nonfurloughed employees on only three items.

Furloughed federal civilian service employees ranked higher in OCB on items two, six, and 13 on the OCB scale. These items were "punctuality," "attendance at work is above the norm," and "make innovative suggestions to improve department." Factors such as age and career experience could have influenced the items on which furloughed employees scored higher than nonfurloughed employees. The sample for nonfurloughed employees contained active duty military members with more diversity in both age and experience than federal civilian service employees who made up the furloughed sample of participants.

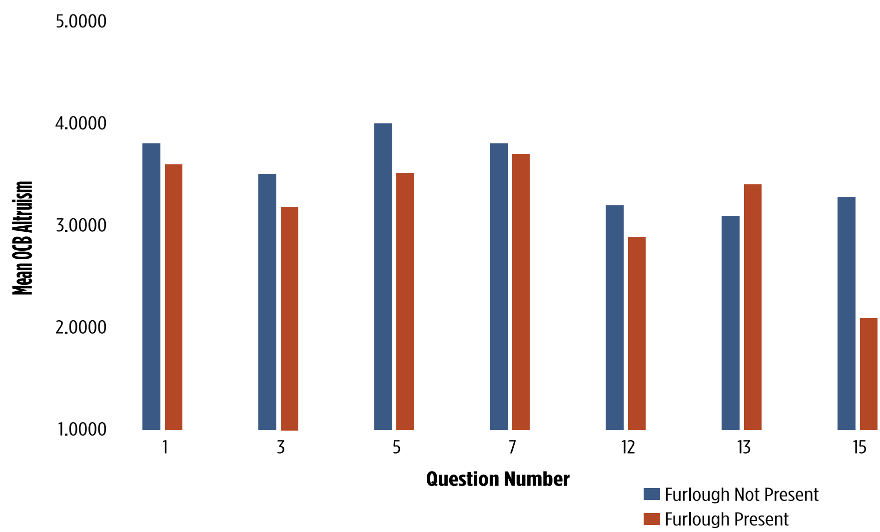
> ***Many U.S. Government federal civilian service employees are retired enlisted active duty military members who tend to be older with more work experience than active duty military members, many of which are just starting their military careers.***

Many U.S. Government federal civilian service employees are retired, enlisted military members. They tend to be older and to have more work experience than active duty military members, many of whom are just starting their military careers. These differences in work experience and age, between furloughed and nonfurloughed employees, could explain why punctuality, attendance, and suggestions to improve the department contradict the comparative scores of the remaining 13 items on the OCB scale.

Organizational Citizenship Behavior: Altruism and General Compliance

The study found no statistically significant difference between furloughed and nonfurloughed employees in the altruism dimension of OCB. The average mean calculated for altruism using the OCB scale for nonfurloughed employees was 3.582 and for furloughed employees was 3.259. Although the results were not statistically significant for altruism, furloughed employees did rank lower in altruism than those not furloughed (Figure 2).

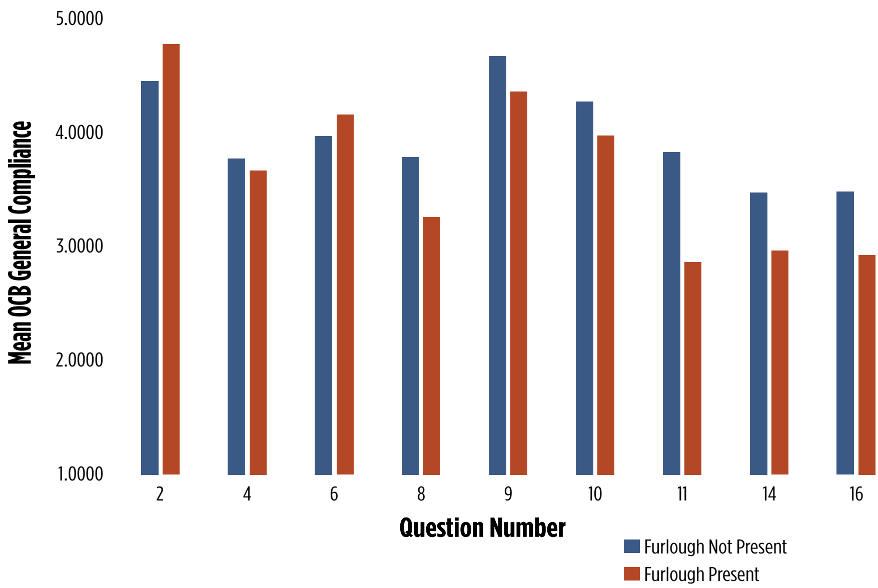
FIGURE 2. MEAN SCORES FOR ORGANIZATIONAL CITIZENSHIP BEHAVIOR: ALTRUISM



1. Help others who have been absent
3. Volunteer for things that are not required
5. Orient new people even though it is not required
7. Help others who have heavy workloads
12. Assist supervisor with his or her work
13. Make innovative suggestions to improve department
15. Attend functions not required but that help the organization image

The general compliance dimension of OCB did reveal a statistically significant difference between furloughed and nonfurloughed federal government employees. Nonfurloughed employees had an average mean of 3.747 for general compliance on the OCB scale, while furloughed employees scored an average mean of 3.674 (Figure 3). Furloughed employees may have lowered their OCB after the furlough period because of an increase in individual workloads. Although furloughed employees were required to take time off without pay during the furlough period, their workload did not change, forcing them to accomplish the same amount of work with less time on the job. It is important to stress that furloughed employees, due to legal implications, were prohibited from working (including telecommuting) during required furlough periods.

FIGURE 3. MEAN SCORES FOR ORGANIZATIONAL CITIZENSHIP BEHAVIOR: GENERAL COMPLIANCE



2. Punctuality
4. Take undeserved breaks (reverse scored)
6. Attendance at work is above the norm
8. Coast toward the end of the day (reverse scored)
9. Give advance notice if unable to come to work
10. Great deal of time spent with personal telephone conversations (reverse scored)
11. Do not take unnecessary time off from work
14. Do not take extra breaks
16. Do not spend time in idle conversation

Implications

The implications of the results of this study contribute to the broader body of knowledge concerning furloughs, OC, and OCB among federal civilian service employees. Very little research had been conducted on furloughs' effects on federal civilian service employees in the areas of OC and OCB. This study finds that furloughs do have an effect on federal civilian service employees and opens the door for future research on this subject.

Because government employees tend to have a greater sense of job security, as pointed out by Kelman (2006), furloughed employees may have resented their furloughs due to a perception of injustice. This study finds that federal civilian service workers who had recently endured a mandated furlough had

lower OC and OCB than those that had not been subjected to a mandated furlough. Without a perception of justification for furloughs among its employees, the federal government could experience higher turnover rates in the future. A high turnover period could prove detrimental to the federal government because the organization relies heavily on the specialized career fields. A large loss of specialized experience within the federal government could make the effect of furloughs counterproductive in the long run, due to an increase in training costs.

> ***Early communication with employees about the potential for furloughs to occur could give employees the opportunity to prepare financially and to adjust their workload in an effort to offset the possibility of stress.***

Recent research has found that it is difficult to change employee perceptions of negatively viewed organizational practice (Totawar & Nambudiri, 2014). Nurturing traits of self-efficacy, optimism, hope, and resilience in employees might ease the negative impact of furloughs on the employees and the organization. The federal government could perhaps reduce the negative impact of furloughs by communicating more effectively. Early communication with employees about the potential for furloughs to occur could give employees the opportunity to prepare financially and to adjust their workload in an effort to offset the possibility of stress. A detailed plan on how the organization intends to remain proficient during furloughs would benefit the organization as well as those employees who must maintain the same standards during furloughs as they would during a period without furloughs. If employees feel that the organization is honest with them, they will be more likely to perceive the actions of the organization as justified.

Limitations

One limitation of this study is the comparison of results collected from civilian employees and active duty military employees. The majority of federal civilian service employees were furloughed at some point during 2013, but active duty military members were not subjected to a furlough during this time. The study does not take into account the two groups' demographic

differences that could have affected the outcome. The study assumes the statistical differences in OC and OCB are the result of the mandated furloughs. Examples of differences between the groups are median age, career experience, and formal job requirements.

The higher OC and OCB of active duty military compared to federal civilian service employees may be because they have more mobility within their formal job requirements. Active duty military have more opportunities to move around to various positions within their organization or to other organizations within the same branch of service when compared to federal civilian service employees. Mobility gives active duty military members the opportunity to seek a good organizational fit, whereas federal civilian service employees lack the same freedom.

Another limitation is that the study methodology does not consider the two groups' internal cultural differences, which exist even though the groups were members of the same organization. The study did not account for the collectivist culture of active duty military members. Differences in the way participants view their obligation to their occupation could have influenced the outcome of this study. Active duty military members may feel a greater sense of commitment to their occupation out of patriotism to their country. Individuals belonging to a collectivist culture tend to take OCB more for granted as a part of their performance than members of a more individualist culture (Lam, Chun, & Law, 1999). The difference between the collectivist culture of active duty military members and the more individualist culture of civilian employees could be another explanation for the disparity of OC and OCB results between the two groups.

A third limitation is the use of a small sample size from a small population within the much larger organization of the federal government. The results of this study are valid based on the sample size for the population of the organization under study. In order for the results to be a true representation of the federal government, however, a much larger sample across multiple organizations within the federal government would have to be surveyed.

A final limitation of this study was the self-reporting methodology used to collect data. Participants rated items based on their perception of their OC and OCB. Participants could exaggerate their level of commitment or OCB for various reasons, such as to make their situation appear worse than it actually is or to appear either more or less committed to their organization than their colleagues. To address this limitation, future researchers could solicit employee managers to complete the surveys based on their employees' behavior. Another option is to conduct a field test on a number of

organizations, utilizing the same survey for the same purpose, in an effort to further validate the use of self-reporting with the OC questionnaire and OCBs scale.

Recommendations for Future Research

Future research conducted on the impact furloughs have on the OC and OCB of federal civilian service employees should take into consideration the limitations of this study. To address those limitations and to improve the understanding of the effect furloughs have on the OC and OCB of federal civilian service employees, this section will provide recommendations for future research. To address the limitation of comparing active duty military members to civilian employees, the researcher recommends a longitudinal study conducted during a furlough period and a period without furloughs. A longitudinal study will enable the researcher to conduct the study on the same organization using only federal civilian service employees. Such a study would eliminate any organizational or occupational differences between divergent groups of employees that could have affected the outcome of this study.

The researcher also recommends that future studies use a larger sample size to ensure external validity. Repeated studies across multiple organizations within the federal government could further validate the findings of the research conducted in this study. A larger study with a similar methodology would provide a much broader understanding of how furloughs affect the OC and OCB of federal government employees without the internal organizational influences of a single organization.

A study of furloughs' effects on organizational factors such as productivity and turnover rates could benefit practitioners in the decision-making process. Because this study found that OC and OCB are negatively affected by furloughs, future research could investigate the relationship between furloughs, OC, OCB, and organizational productivity. With more information on how furloughs affect employee productivity and turnover rates, administrators will be able to make more educated decisions on how to implement furloughs while minimizing loss of organizational assets.

More research conducted on mandated furloughs utilizing federal civilian service employees could prove useful to both scholars and practitioners due to unique qualities in career fields, job requirements, employee stressors, and organizational culture. Future research in the area of furloughs in the federal government could reveal more effective methods to reduce the employee stress and anxiety brought on by feelings of uncertainty in the wake of furloughs. A final recommendation for future research is

to investigate ways to reduce the negative effects of furloughs on federal civilian service employees. One avenue for future research could be to investigate the relationship between positive reinforcement, furloughs, OC, and OCB.

Conclusion

The results of this study show that furloughs have a negative effect on the OC and OCB of U.S. Government federal civilian service employees. A reduction in OC and OCB due to furloughs could lead to lower production and higher turnover rates, resulting in higher expenditures. Because the U.S. Government has had very few substantial furloughs in the recent past, there is very little research on how U.S. Government federal civilian service employees react to a sustained mandated furlough.

Managing the effects of furloughs within the federal government is an exceptional challenge for both supervisors and their employees because most furloughs are unplanned due to a failure to pass a budget or continuing resolution. Although unpredictability is inevitable, supervisors and employees can take steps to minimize the stress caused by both planned and unplanned furloughs. Supervisors can encourage employees to plan for furloughs financially, and they can implement open door policies to address concerns about the stressors accompanying furloughs, such as increased workloads before and after the furlough. Administrators within the federal government should continue to gather information in an effort to understand how to implement furloughs successfully with minimum strain on organizations and employees. It is in the best interest of the U.S. Government to proactively retain a productive and dedicated workforce while finding ways to meet the financial challenges that lie ahead.

References

Ariani, D. W. (2012). Leader-member exchanges as a mediator of the effect of job satisfaction on affective organizational commitment: An empirical test. *International Journal of Management, 29*(1), 46–56.

Badiru, A. (2014). Adverse impacts of furlough programs on employee work rate and organizational productivity. *Defense Acquisition Research Journal, 21*(2), 595–624.

Bolino, M. C., Klotz, A. C., Turnley, W. H., & Harvey, J. (2013). Exploring the dark side of organizational citizenship behavior. *Journal of Organizational Behavior, 34*(4), 542–559.

Brutocao, S., & Marshall, A. N. (2011). Cost-saving options for employers and other wage and hour issues. *ABA Journal of Labor & Employment Law, 26*(3), 383–404.

Coyle-Shapiro, J. A. M., & Conway, N. (2005). Exchange relationships: Examining psychological contracts and perceived organizational support. *Journal of Applied Psychology, 90*(4), 774–781.

Cristian-Liviu, V. (2013). Organizational culture and strategy. How does it work? An empirical research. *Annals of the University of Oradea, Economic Science Series, 22*(1), 1690–1696.

DelCampo, R. (2007). Psychological contract violations: An individual difference perspective. *International Journal of Management, 24*(1), 43–52.

Duggan, M. J., Lewis, J. E., & Milluzzi, M. (2010). Labor relations in hard times. *Illinois Public Employee Relations Report, 27*(2), 1–8.

Elias, S. M. (2007). Employee commitment in times of change: Assessing the importance of attitudes toward organizational change. *Journal of Management, 35*(1), 37–55.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

Ferrell, B. S. (2014). *Sequestration: Comprehensive and updated cost savings would better inform DOD decision makers if future civilian furloughs occur* (GAO-14-529). Retrieved from http://www.gao.gov/assets/670/664142.pdf

Gurbuz, S. (2009). Some possible antecedents of military personnel organizational citizenship behavior. *Military Psychology, 21*(2), 200–215.

Halbesleben, J. B., Wheeler, A. R., & Paustian-Underdahl, S. C. (2013). The impact of furloughs on emotional exhaustion, self-rated performance, and recovery experiences. *Journal of Applied Psychology, 98*(3), 492–503.

Hohman, M., Packard, T., Finnegan, D., & Jones L. (2013). The effect of mandatory furloughs on self-determination, financial strain, and decision to leave the California State University system in social work faculty. *Journal of Social Work Education, 49*, 748–759.

Kelman, S. (2006). Downsizing, competition, and organizational change in government: Is necessity the mother of invention? *Journal of Policy Analysis & Management, 25*(4), 875–895.

Lam, S. K., Chun, H., & Law, K. S. (1999). Organizational citizenship behavior: Comparing perspectives of supervisors and subordinates across four international samples. *Journal of Applied Psychology, 84*(4), 594–601.

Lee, S., & Sanders, R. M. (2013). Fridays are furlough days: The impact of furlough policy and strategies for human resource management during a severe economic recession. *Review of Public Personnel Administration, 33*(3), 299–311.
Osborne, D., Smith, H., & Huo, Y. (2013). More than a feeling: Discrete emotions mediate the relationship between relative deprivation and reaction to workplace furloughs. *Personality and Social Psychology Bulletin, 38*(5), 628–641.
Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior, 14*(2), 224–227.
Nadin, S., & Cassell, C. (2007). New deal for old? Exploring the psychological contract in a small firm environment. *International Small Business Journal, 25*(4), 417–443.
Shahnawaz, M. G., & Goswami, K. (2011). Effect of psychological contract violation on organizational commitment, trust and turnover intention in private and public sector Indian organizations. *Vision: The Journal of Business Perspective, 15*(3), 209–217.
Shannon, R. (2010). Public sector furloughs: Player perspectives, strategies, and grounds for challenge. *Illinois Public Employee Relations Report, 27*(4), 1–9.
Smith, C., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology, 68*(4), 653–663.
Totawar, A., & Nambudiri, R. (2014). How does organizational justice influence job satisfaction and organizational commitment? Explaining with psychological capital. *Vikalpa: The Journal for Decision Makers, 39*(2), 83–97.

Biography

Dr. Robert L. Shepherd is employed as an airfield systems maintenance technician at Eglin Air Force Base in Florida. He recently graduated from Capella University with a PhD in Organizational Management and Leadership. His dissertation concentrated on the effect that federally mandated furloughs have on the commitment of federal civilian service workers. He graduated with a master's degree in Organizational Management from Troy University in 2008.

(E-mail address: robert.shepherd.2@us.af.mil)

RECOGNITION OF REVIEWERS
FOR PRINT YEAR 2016

We would like to express our appreciation to all of the subject matter experts who volunteered to participate in the *Defense Acquisition Research Journal* peer review process. The assistance of these individuals provided impartial evaluation of the articles published during the 2016 print year. We would also like to acknowledge those referees who wished to remain anonymous. Your continued support is greatly appreciated, and we look forward to working with many of you again in print year 2017.

Dr. Craig M. Arndt
Senior Executive Fellow
Professor of Systems Engineering Management
Defense Acquisition University

Dr. Michelle Bailey
Professor of Engineering
Defense Acquisition University

Mr. William A. Broadus III
Professor of Engineering and Program Management
Defense Acquisition University

Mrs. Penny Brooks
Professor of Contract Management
Defense Acquisition University

Ms. Clareice Chaney
Principal Information Systems Engineer
The MITRE Corporation

Dr. John M. Colombi
Associate Professor of Systems Engineering
Air Force Institute of Technology

COL John T Dillard, USA (Ret.)
Academic Associate for Systems Acquisition Management
Naval Postgraduate School

Mr. David H. Dotson
Professor of Contract Management
Defense Acquisition University

Mr. Salvatore Fazio
Professor of Acquisition Management
Defense Acquisition University

Ms. Georgella McRae Fordham
Professor of Contract Management
Defense Acquisition University

RECOGNITION OF REVIEWERS (CONT.)

Mr. Donald E. Goddard
Professor of Systems Engineering
Defense Acquisition University

Dr. Rebecca A. Grier
Research Staff Member
Institute for Defense Analyses

Mr. Raymond D. Jones
Lecturer
Naval Postgraduate School

Mr. Jay N. Kanavos
Acquisition Lead, Center for Acquisition and Management Sciences
The MITRE Corporation

Mr. John W. Kelley
Professor of Acquisition Management
Defense Acquisition University

Dr. Clay Koschnick
Assistant Professor of Systems Engineering
Air Force Institute of Technology

Lt Col Brandon M. Lucas, USAF, Ph.D.
Director, Graduate Cost Analysis Program
Air Force Institute of Technology

Ms. Janie L. Maddox
Lecturer
Naval Postgraduate School

Dr. Paul J. Mallon
Professor of Program Management
Defense Acquisition University

Mr. Timothy G. Mead
Professor of Engineering and Technology
Defense Acquisition University

RECOGNITION OF REVIEWERS (CONT.)

Mr. Patrick Morrow
Professor of Systems Acquisition Management
Defense Acquisition University

COL Robert F. Mortlock, USA (Ret.), Ph.D.
Lecturer
Naval Postgraduate School

Mr. Nicholas Orlans
Senior Principal Engineer
The MITRE Corporation

Lt Col Stephen R. Parks, USAF
Professor of Acquisition Management
Defense Acquisition University

Mr. Thomas R. Rauls
Professor of Acquisition Management
Defense Acquisition University

Mr. David M. Riel
Professor of Acquisition Management
Defense Acquisition University

Dr. Susan L. Rose
Research Staff Member
Institute for Defense Analyses

Dr. Stanley G. Rosen
Professor of Acquisition Management
Defense Acquisition University

Mr. Brian E. Schultz
Professor of Program Management
Defense Acquisition University

Mr. Jonathan Specht
Professor of Systems Engineering
Defense Acquisition University

Mrs. Dana Stewart
Professor of Financial Management
Defense Acquisition University

Professor Trevor Taylor
Emeritus Professor
Cranfield University

Dr. Richard Volkert
Lead Systems Engineer ISR/IO Department
SSC Pacific

RECOGNITION OF REVIEWERS (CONT.)

Dr. Charlette W. Watkins
Principal, Program Management
Acquisition
The MITRE Corporation

Mr. Michael S. Wilson
Lead Business Analyst
The MITRE Corporation

Ms. Virginia L. Wydler
Principal Analyst
The MITRE Corporation

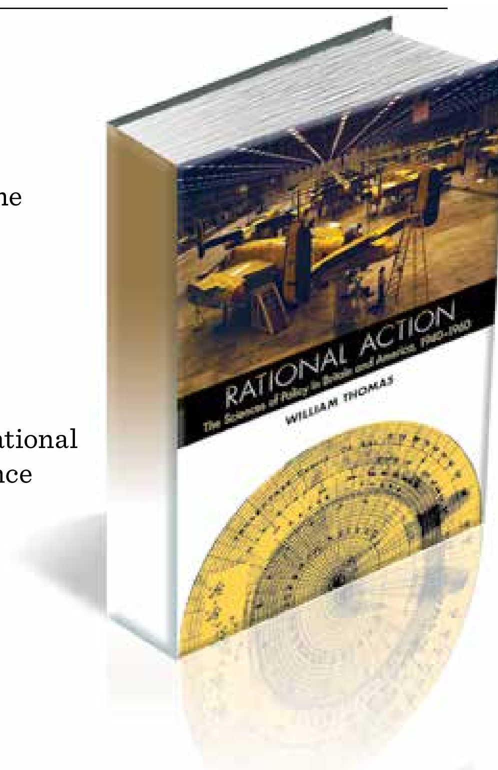
PROFESSIONAL READING LIST

The Defense Acquisition Professional Reading List is intended to enrich the knowledge and understanding of the civilian, military, contractor, and industrial workforce who participate in the entire defense acquisition enterprise. These book recommendations are designed to complement the education and training vital to developing essential competencies and skills of the acquisition workforce. Each issue of the *Defense Acquisition Research Journal* will include one or more reviews of suggested books, with more available on our Website http://dau.dodlive.mil/defense-acquisition-professional-reading-program.

We encourage our readers to submit book reviews they believe should be required reading for the defense acquisition professional. The books themselves should be in print or generally available to a wide audience; address subjects and themes that have broad applicability to defense acquisition professionals; and provide context for the reader, not prescriptive practices. Book reviews should be 450 words or fewer, describe the book and its major ideas, and explain its relevancy to defense acquisition. Please send your reviews to the managing editor, *Defense Acquisition Research Journal* at *DefenseARJ@dau.mil*.

Featured Book

Rational Action: The Sciences of Policy in Britain and America, 1940–1960
Series: Transformations: Studies in the History of Science and Technology
Author: William Thomas
Publisher: The MIT Press
Copyright Date: 2015
Hardcover: 416 pages
ISBN: 9780262028509
Reviewed by: Petros Boutselis, Operational Researcher/Lecturer, Centre for Defence Acquisition, Cranfield University, UK



Review:

It was Plato who said, “Necessity is the mother of invention,” and there cannot be a more invention-driven context than the one imposed by a war. The foundations of Operational Research (OR) were laid during the world wars, and today OR is both a science and a practice that has been incorporated in defense and business activities since then. William Thomas takes readers on a journey through the stages of its development, offering insight to the problems that necessitated its birth. These problems are not very different from what today’s defense decision makers manage, ranging from the type of systems with which a unit should be armed and the required support, to how to deploy and operate these units.

However, decisions today that are related to units’ acquisition and support are more frequent, and in that sense also more intense, than decisions on their operational use. *Rational Action* does not offer the textbook list of best practices to such problems and, indeed, decision makers need more than a guide denoting what procedures to follow. They also need the insights obtained by the “what happened” question in order to relate to similar cases and understand how they evolved. Social and political factors cannot be joined with the practical reality of calling for action to arrive at solutions that are both feasible and desirable, except by being viewed through the lens offered by credible historical research. William Thomas offers just that.

“What [Leslie Bennett Craigie] Cunningham’s theory did [in 1937 for the configuration of aircraft design] was probabilistically interrelate armament specifications with mathematical representations of air combat scenarios” (p. 34), which closely resembles how specifications are still developed today. Further, on p. 35 we follow within the historical narrative the existence of the unavoidable political drivers: “[Cunningham’s] paper was so fortunate to survive the [Royal Aircraft Establishment’s] serious fundamental criticisms again …,” and on p. 103: “[In 1940, Warren Weaver] showed and praised [Cunningham’s papers] to various military officers until they asked him to try to digest and simplify them, and to explain their content in terms not so formidably mathematical.”

In an even larger acquisition challenge—that of 1944’s B-29 “Superfortress” bomber—William Thomas offers similar significant insights. Warren Weaver’s Applied Mathematical Panel took over the AC-92 contract to improve B-29 bombing accuracy (p. 121) while “Unfortunately, there was very little theory, test data, manuals, or training standards ….” The continuation is equally enlightening: “Throughout the course of the AC-92 contract,

Weaver steadily and stubbornly insisted that all aspects of the problem needed to be worked ... A sense of the theoretical and empirical disarray in which work on AC-92 took place can be gained from an internal working paper on the question of whether the B-29 should be modified ...” (p. 122).

Similar work has been done to the support dimension of acquired systems. Thomas, in his chapter on the development of “Theories of Decision, Allocation and Design” (chapter 20), discusses the Travelling Salesman Problem, and he devotes chapter 21 to Inventory Theory.

Even though *Rational Action* addresses matters of the past, it is of interest to historians or operational researchers whose science foundations are the subject of the narratives. Through the historical analysis of obsolete systems and already-solved problems, it becomes essential reading for managers and decision makers who want a better understanding of the critical factors that drove the evolution of acquisition-related problems.

Statement Required by the Act of August 12, 1970 Section 3685, Title 39, U.S.C. Showing Ownership, Management, and Circulation

The ***Defense Acquisition Research Journal*** *(ARJ)* is published at the Defense Acquisition University (DAU), Fort Belvoir, VA. The University publishes four issues annually. The Director of the Defense Acquisition University (DAU) Press is Randy Weekes, the Managing Editor of the *Defense ARJ* is Norene Taylor, and the publisher is the DAU Press. All are collocated at the following address:

Defense Acquisition University
Attn DAU Press (Defense ARJ)
9820 Belvoir Road, Suite 3
Fort Belvoir, VA 22060-5565

Average Number of Copies Each Issue During the Preceding 12 Months:

A.	Total number of copies printed (net press run)	3,343
B.	Paid and/or requested circulation:	
	1. Sales through dealers and carriers, street vendors, and counter sales	0
	2. Mail subscriptions paid and/or requested	3,189
C.	Total paid and/or requested circulation	3,189
D.	Free distribution by mail, carrier, or other means; samples, complimentary, and other free copies	133
E.	Total distribution	3,322
F.	Copies not distributed:	
	1. Office use, leftover, unaccounted, spoiled after printing	21
	2. Returns from news agents	0
G.	Total	3,343

Actual Number of Copies of Single Issue Published Nearest to Filing Date:

A.	Total number of copies printed (net press run)	3,341
B.	Paid and/or requested circulation:	
	1. Sales through dealers and carriers, street vendors, and counter sales	0
	2. Mail subscriptions paid and/or requested	3,176
C.	Total paid and/or requested circulation	3,176
D.	Free distribution by mail, carrier, or other means; samples, complementary, and other free copies	134
E.	Total distribution	3,310
F.	Copies not distributed:	
	1. Office use, leftover, unaccounted, spoiled after printing	31
	2. Returns from news agents	0
G.	Total	3,341

Defense ARJ and *AT&L* have become online-only publications for individual subscribers

If you would like to start or continue a membership with Defense ARJ or AT&L, you must register a valid e-mail address in our **LISTSERV**

NEW

Online presence for easier use on mobile and desktop devices: **http://dau.dodlive.mil**

Printable PDF files can be found at **http://www.dau.mil/publications**

ALL READERS: Please subscribe or resubscribe so you will not miss out on receiving future publications.

- Send an e-mail to **darjonline@dau.mil** and/or **datlonline@dau.mil**, giving the e-mail address you want us to use to notify you when a new issue is posted.
- Please type “Add to LISTSERV” in the subject line.
- Please also use this address to notify us if you change your e-mail address.

Defense ARJ Guidelines FOR CONTRIBUTORS

The *Defense Acquisition Research Journal (ARJ)* is a scholarly peer-reviewed journal published by the Defense Acquisition University (DAU). All submissions receive a blind review to ensure impartial evaluation.

IN GENERAL

We welcome submissions from anyone involved in the defense acquisition process. *Defense acquisition* is defined as the conceptualization, initiation, design, development, testing, contracting, production, deployment, logistics support, modification, and disposal of weapons and other systems, supplies, or services needed for a nation's defense and security, or intended for use to support military missions.

Research involves the creation of new knowledge. This generally requires using material from primary sources, including program documents, policy papers, memoranda, surveys, interviews, etc. Articles are characterized by a systematic inquiry into a subject to discover/revise facts or theories with the possibility of influencing the development of acquisition policy and/or process.

We encourage prospective writers to coauthor, adding depth to manuscripts. It is recommended that a mentor be selected who has been previously published or has expertise in the manuscript's subject. Authors should be familiar with the style and format of previous *Defense ARJs* and adhere to the use of endnotes versus footnotes (refrain from using the electronic embedding of footnotes), formatting of reference lists, and the use of designated style guides. *It is also the responsibility of the corresponding author to furnish any required government agency/employer clearances with each submission.*

MANUSCRIPTS

Manuscripts should reflect research of empirically supported experience in one or more of the areas of acquisition discussed above. Empirical research findings are based on acquired knowledge and experience versus results founded on theory and belief. Critical characteristics of empirical research articles:

- clearly state the question,
- define the methodology,
- describe the research instrument,
- describe the limitations of the research,
- ensure results are quantitative and qualitative,
- determine if the study can be replicated, and
- discuss suggestions for future research (if applicable).

Research articles may be published either in print and online, or as a Web-only version. Articles that are 4,500 words or less (excluding abstracts, references, and endnotes) will be considered for print as well as Web publication. Articles between 4,500 and 10,000 words will be considered for Web-only publication, with an abstract (150 words or less) included in the print version of the *Defense ARJ*. In no case should article submissions exceed 10,000 words.

Book Reviews

Defense ARJ readers are encouraged to submit reviews of books they believe should be required reading for the defense acquisition professional. The reviews should be 450 words or fewer describing the book and its major ideas, and explaining why it is relevant to defense acquisition. In general, book reviews should reflect specific in-depth knowledge and understanding that is uniquely applicable to the acquisition and life cycle of large complex defense systems and services.

Audience and Writing Style

The readers of the *Defense ARJ* are primarily practitioners within the defense acquisition community. Authors should therefore strive to demonstrate, clearly and concisely, how their work affects this community. At the same time, do not take an overly scholarly approach in either content or language.

Format

Please submit your manuscript with references in APA format (author-date-page number form of citation) as outlined in the *Publication Manual of the American Psychological Association* (6th Edition). For all other style questions, please refer to the *Chicago Manual of Style* (16th Edition).

Contributors are encouraged to seek the advice of a reference librarian in completing citation of government documents because standard formulas of citations may provide incomplete information in reference to government works. Helpful guidance is also available in *The Complete Guide to Citing Government Documents* (Revised Edition): *A Manual for Writers and Librarians* (Garner & Smith, 1993), Bethesda, MD: Congressional Information Service.

Pages should be double-spaced in Microsoft Word format, Times New Roman, 12-point font size and organized in the following order: title page (titles, 12 words or less), abstract (150 words or less to conform with formatting and layout requirements of the publication), two-line summary, list of keywords (five words or less), reference list (only include works cited in the paper), author's note or acknowledgments (if applicable), and figures or tables (if any). Manuscripts submitted as PDFs will not be accepted.

Figures or tables should not be inserted or embedded into the text, but segregated (one to a page) at the end of the text. When material is submitted electronically, *each figure or table should be saved to a separate, exportable file* (i.e., a readable EPS file). For additional information on the preparation of figures or tables, refer to the Scientific Illustration Committee, 1988,

Illustrating Science: Standards for Publication, Bethesda, MD: Council of Biology Editors, Inc. Restructure briefing charts and slides to look similar to those in previous issues of the *Defense ARJ*.

The author (or corresponding author in cases of multiple authors) *should attach a signed cover letter* to the manuscript that provides all of the authors' names, mailing and e-mail addresses, as well as telephone and fax numbers. The letter should verify that the submission is an original product of the author(s); that all the named authors materially contributed to the research and writing of the paper; that the submission has not been previously published in another journal (monographs and conference proceedings serve as exceptions to this policy and are eligible for consideration for publication in the *Defense ARJ*); and that it is not under consideration by another journal for publication. Details about the manuscript should also be included in the cover letter: for example, title, word length, a description of the computer application programs, and file names used on enclosed DVD/CDs, e-mail attachments, or other electronic media.

COPYRIGHT

The *Defense ARJ* is a publication of the United States Government and as such is not copyrighted. Because the *Defense ARJ* is posted as a complete document on the DAU homepage, we will not accept copyrighted manuscripts that require special posting requirements or restrictions. If we do publish your copyrighted article, we will print only the usual caveats. The work of federal employees undertaken as part of their official duties is not subject to copyright except in rare cases.

Web-only publications will be held to the same high standards and scrutiny as articles that appear in the printed version of the journal and will be posted to the DAU Website at www.dau.mil.

In citing the work of others, please be precise when following the author-date-page number format. It is the contributor's responsibility to obtain permission from a copyright holder if the proposed use exceeds the fair use provisions of the law (see U.S. Government Printing Office, 1994, *Circular 92: Copyright Law of the United States of America*, p. 15, Washington, DC). Contributors will be required to submit a copy of the writer's permission to the managing editor before publication.

We reserve the right to decline any article that fails to meet the following copyright requirements:

- The author cannot obtain permission to use previously copyrighted material (e.g., graphs or illustrations) in the article.
- The author will not allow DAU to post the article in our *Defense ARJ* issue on our Internet homepage.
- The author requires that usual copyright notices be posted with the article.
- To publish the article requires copyright payment by the DAU Press.

SUBMISSION

All manuscript submissions should include the following:

- Cover letter
- Author checklist
- Biographical sketch for each author (70 words or less)
- Headshot for each author should be saved to a CD-R disk or e-mailed at 300 dpi (dots per inch) or as a high-print quality JPEG or Tiff file saved at no less than 5x7 with a plain background in business dress for men (shirt, tie, and jacket) and business appropriate attire for women. All active duty military should submit headshots in Class A uniforms. Please note: low-resolution images from Web, Microsoft PowerPoint, or Word will not be accepted due to low image quality.
- One copy of the typed manuscript, including:
    - Title (12 words or less)
    - Abstract of article (150 words or less)
    - Two-line summary
    - Keywords (5 words or less)
    - Document double-spaced in Microsoft Word format, Times New Roman, 12-point font size (4,500 words or less for the printed edition and 10,000 words or less for the online-only content excluding abstract, figures, tables, and references).

These items should be sent electronically, as appropriately labeled files, to the *Defense ARJ* Managing Editor at: DefenseARJ@dau.mil.

CALL FOR AUTHORS

We are currently soliciting articles and subject matter experts for the 2017 *Defense Acquisition Research Journal* (*ARJ*) print year. Please see our guidelines for contributors for submission deadlines.

Even if your agency does not require you to publish, consider these career-enhancing possibilities:

- Share your acquisition research results with the Acquisition, Technology, and Logistics (AT&L) community.
- Change the way Department of Defense (DoD) does business.
- Help others avoid pitfalls with lessons learned or best practices from your project or program.
- Teach others with a step-by-step tutorial on a process or approach.
- Share new information that your program has uncovered or discovered through the implementation of new initiatives.
- Condense your graduate project into something beneficial to acquisition professionals.

ENJOY THESE BENEFITS:

- Earn 25 continuous learning points for publishing in a refereed journal.
- Earn a promotion or an award.
- Become part of a focus group sharing similar interests.
- Become a nationally recognized expert in your field or specialty.
- Be asked to speak at a conference or symposium.

We welcome submissions from anyone involved with or interested in the defense acquisition process—the conceptualization, initiation, design, testing, contracting, production, deployment, logistics support, modification, and disposal of weapons and other systems, supplies, or services (including construction) needed by the DoD, or intended for use to support military missions.

If you are interested, contact the Defense ARJ *managing editor (**DefenseARJ@dau.mil**) and provide contact information and a brief description of your article. Please visit the* Defense ARJ *Guidelines for Contributors at **http://www.dau.mil/publications/DefenseARJ/ARJ/ARJ79/ARJ79-GuidelinesContributors.pdf**.*

Defense ARJ PRINT SCHEDULE

The *Defense ARJ* is published in quarterly theme editions. All submissions are due by the first day of the month. See print schedule below.

Author Deadline	Issue
July	January
November	April
January	July
April	October

In most cases, the author will be notified that the submission has been received within 48 hours of its arrival. Following an initial review, submissions will be referred to peer reviewers and for subsequent consideration by the Executive Editor, *Defense ARJ*.

Contributors may direct their questions to the Managing Editor, *Defense ARJ*, at the address shown below, or by calling 703-805-3801 (fax: 703-805-2917), or via the Internet at norene.taylor@dau.mil.

The DAU Homepage can be accessed at:
http://www.dau.mil

DEPARTMENT OF DEFENSE

DEFENSE ACQUISITION UNIVERSITY

ATTN: DAU PRESS (*Defense ARJ*)

9820 BELVOIR RD STE 3

FORT BELVOIR, VA 22060-5565


Now you can search the DAU Website
and our online publications!
Defense Acquisition University
WEBSITES
http://www.dau.mil
and
http://dau.dodlive.mil
Your Online Access to Acquisition Research, Consulting,
Information, and Course Offerings

ARJ

Defense Acquisition Research Journal

SURVEY

Please rate this publication based on the following scores:

5 —Exceptional 4 — Great 3 — Good 2 — Fair 1 — Poor

Please circle the appropriate response.

1. How would you rate the overall publication?	5	4	3	2	1
2. How would you rate the design of the publication?	5	4	3	2	1

	True	False
a) This publication is easy to read		
b) This publication is useful to my career		
c) This publication contributes to my job effectiveness		
d) I read most of this publication		
e) I recommend this publication to others in the acquisition field		

If hand written, please write legibly.

3. What topics would you like to see get more coverage in future *Defense ARJs*?

4. What topics would you like to see get less coverage in future *Defense ARJs*?

5. Provide any constructive criticism to help us to improve this publication:

6. Please provide e-mail address for follow up (optional):

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 12 FORT BELVOIR,VA

POSTAGE WILL BE PAID BY ADDRESSEE

DEFENSE ACQUISITION UNIVERSITY
ATTN: DAU PRESS
9820 BELVOIR RD STE 3
FORT BELVOIR VA 22060-9910

FREE ONLINE SUBSCRIPTION

☐ *Defense ARJ* *Defense AT&L* ☐

Thank you for your interest in *Defense Acquisition Research Journal* and *Defense AT&L* magazine. To receive your complimentary online subscription, please write legibly if hand written and answer all questions below—incomplete forms cannot be processed.

***When registering, please do not include your rank, grade, service, or other personal identifiers.**

☐ *New Online Subscription*

☐ *Cancellation*

☐ *Change E-mail Address*

Date

Last Name:

First Name:

Day/Work Phone:

E-mail Address:

Signature: *(Required)*

PLEASE FAX TO: 703-805-2917

The Privacy Act and Freedom of Information Act

In accordance with the Privacy Act and Freedom of Information Act, we will only contact you regarding your *Defense ARJ* and *Defense AT&L* subscriptions. If you provide us with your business e-mail address, you may become part of a mailing list we are required to provide to other agencies who request the lists as public information. If you prefer not to be part of these lists, please use your personal e-mail address.

SUBSCRIPTION

ver 01/03/2017

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 12 FORT BELVOIR,VA

POSTAGE WILL BE PAID BY ADDRESSEE

DEFENSE ACQUISITION UNIVERSITY
ATTN: DAU PRESS
9820 BELVOIR RD STE 3
FORT BELVOIR VA 22060-9910

ARJ

Defense Acquisition Research Journal

A Publication of the Defense Acquisition University

ARJ

We're on the Web at:
http://www.dau.mil/publications/DefenseARJ
and
http://dau.dodlive.mil

ARJ

Defense Acquisition Research Journal

A Publication of the Defense Acquisition University

Learn. Perform. Succeed.

WINNER 2015 APEX® AWARDS FOR PUBLICATION EXCELLENCE

WINNER 2016 APEX® AWARDS FOR PUBLICATION EXCELLENCE